Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets

Gilles de Truchis, Florent Dubois

Working Papers / Documents de travail

Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets

Gilles de Truchis
Florent Dubois

WP 2014 - Nr 45

Unbalanced Fractional Cointegration and the No-Arbitrage Condition on Commodity Markets

Gilles de Truchis[1,*], Florent Dubois[1]

Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

Preliminary draft
September 2014

Technical and non-technical abstracts

A necessary condition for two time series to be nontrivially cointegrated is the equality of their respective integration orders. Nonetheless, in some cases, the apparent unbalance of integration orders of the observables can be misleading and the cointegration theory applies all the same. This situation refers to unbalanced cointegration in the sense that balanced long run relationship can be recovered by an appropriate filtering of one of the time series. In this paper, we suggest a local Whittle estimator of bivariate unbalanced fractional cointegration systems. Focusing on a degenerating band around the origin, it estimates jointly the unbalance parameter, the long run coefficient and the integration orders of the regressor and the cointegrating errors. Its consistency is demonstrated for the stationary regions of the parameter space and a finite sample analysis is conducted by means of Monte Carlo experiments. An application to the no-arbitrage condition between crude oil spot and futures prices is proposed to illustrate the empirical relevance of the developed estimator.

The no-arbitrage condition between spot and future prices implies an analogous condition on their underlying volatilities. Interestingly, the long memory behavior of the volatility series also involves a long-run relationship that allows to test for the no-arbitrage condition by means of cointegration techniques. Unfortunately, the persistent nature of the volatility can vary with the future maturity, thereby leading to unbalanced integration orders between spot and future volatility series. Nonetheless, if a balanced long-run relationship can be recovered by an appropriate filtering of one of the time series, the cointegration theory applies all the same and unbalanced cointegration operates between the raw series. In this paper, we introduce a new estimator of unbalanced fractional cointegration systems that allows to test for the no-arbitrage condition between the crude oil spot and futures volatilities.

Keywords: Unbalanced cointegration, Fractional cointegration, No-arbitrage condition, Local Whittle likelihood, Commodity markets
JEL: C22, G10

---

1. Introduction

This paper deals with the estimation of a general class of models, terms unbalanced cointegration systems, that encompass the well known triangular cointegration system. The former systems originate from very recent developments (see Hualde 2006), while the latter relies on the seminal definition of

---

[*]Corresponding author

[**]We are grateful to Guillaume Chevillon, Mathieu Faure, Sébastien Laurent and Vêlayoudom Marimoutou for helpful comments and discussions.

Email addresses: `gilles.de-truchis@univ-amu.fr` (Gilles de Truchis), `florent.dubois@univ-amu.fr` (Florent Dubois)

[1]Aix-Marseille School of Economics, Aix-Marseille University, Château Lafarge, Route des Milles 13290 Aix-en-Provence, France. Tel.: 00 33 (0)4 42 93 59 60; fax: 00 33 (0)4 42 38 95 85.

the cointegration (see Granger 1981) which establishes that two time series y_t and x_t share a common stochastic trend, if (i) y_t and x_t are both integrated of order δ_2, hereafter $I(\delta_2)$; (ii) there exists a non-null scalar β so that $e_t = y_t - \beta x_t \sim I(\delta_1)$ and $\delta_2 - \delta_1 > 0$.[2] This definition is quite general and does not constrain integration orders to be integers. Nonetheless, following the paper of Engle and Granger (1987), the literature has primarily investigated the particular case where observables are unit root processes, i.e. $I(1)$, and a linear combination between them has short memory, i.e. $I(0)$.

In a pioneer work, Cheung and Lai (1993) extended the model of Engle and Granger (1987), allowing the integration order of the cointegration errors (i.e. δ_1) to be a real number. Their methodology is simple and operates in two steps if y_t and x_t are $I(1)$. When y_t and x_t are fractionally integrated (i.e. $\delta_2 \in \mathbb{R}$), some new complications arise. While a part of the literature has focused on testing the homogeneity of integration orders between y_t and x_t (see e.g Robinson and Yajima 2002, Nielsen and Shimotsu 2007, Hualde 2013), another part has investigated the estimation of such systems with unknown integration orders (see e.g Robinson and Hualde 2003). In the latter case, many difficulties appear in the uniform treatment of the objective functions on the entire parameter space. Thereby, three cases are generally distinguished in the literature. Considering that a fractionally integrated process, is stationary when the integration order is less than $1/2$ and nonstationary otherwise, and interpreting $\delta_2 - \delta_1$ as the cointegration strength, (i) the strong cointegration occurs when $\delta_2 - \delta_1 > 1/2$; (ii) the weak cointegration occurs when $\delta_2 - \delta_1 < 1/2$; (iii) the stationary cointegration occurs when $\delta_1 < \delta_2 < 1/2$. As Velasco (2003) suggested, it is more efficient to estimate simultaneously all the parameters of interest (see also Lobato 1999). Among recent contributions in this direction we can mention Hualde and Robinson (2007) and Shimotsu (2012) concerning the weak and strong cointegration cases and Nielsen (2007) and Robinson (2008) for the stationary case.[3] In this paper we are particularly interested in the stationary cointegration case which is mainly attractive in empirical finance where time series have long range dependence but are likely to be stationary (e.g. volatility, volume, closing prices of commodities). Investigating stationary cointegration is also of interest because spurious regression can occurs, whether y_t and x_t are stationary or not, as long as their integration orders sum up to a value greater than $1/2$ (see Tsay and Chung 2000). Nonetheless, it is not so easy to identify whether or not the observed time series are stationary leading practitioners to possibly make a wrong decision. Accordingly, we are also interested in non-stationary region of the parameter space, although the proposed estimator is not theoretically designed to handle this case. In the present paper, this issue is investigating by means of Monte Carlo study.

Meanwhile, Hualde (2006) investigated a promising alternative avenue of research termed unbalanced cointegration. Let y_t and x_t be two observable time series integrated of orders δ_2 and $\delta_2 + \xi$ respectively. Following Hualde (2006), unbalanced cointegration is likely to occur between y_t and x_t and cointegration theory, in the usual sense, is likely to apply between y_t and $x_t(\xi)$, if there exists a linear combination

[2]See Granger (2010) for a very simple and short introduction to the cointegration theory.

[3]See also Dueker and Startz (1998) and de Truchis (2013) for a fully parametric approach in time and frequency domain respectively. Hualde and Robinson (2010) and Johansen and Nielsen (2012) deal with extensions to multivariate case in frequency and time domain respectively, but this discussion goes beyond of the scope of the paper.

between them which has less memory δ_1. Accordingly, this type of cointegration does not differ from the original cointegration theory but is useful from an empirical point of view. Hualde (2006) suggests to estimate ξ by the difference between integration orders of y_t and x_t and discusses the consistency of the OLS estimator of β. With respect to the joint estimation of β and ξ, the standard OLS are unfeasible and Hualde (2014) investigated the consistency of the non-linear least squares estimates of ξ and β when $\delta_2 > 1/2$ and $\delta_1 \geq 0$.[4] Conversely, we focus here on the stationary case assuming that $\delta_2 \in (0, 1/2 - |\xi|)$ and $\delta_2 > \delta_1 \geq 0$. We propose a local Whittle estimator of bivariate unbalanced fractional cointegration systems. Focusing on a degenerating band around the origin of the spectral density matrix, it estimates jointly the unbalance parameter, the long run coefficient and the integration orders of the regressor and the cointegrating errors without specifying the short run dynamics avoiding thereby the misspecification issue. The consistency of the estimator is discussed and highlight that β is $(m/n)^{\delta_1-\delta_2}$-consistent when x_t is appropriately filtered in the long run equation, with m the bandwidth number and n the sample size. The finite sample properties are also investigated by means of Monte Carlo simulations for a wide range of specifications.

In the empirical part of the paper, we aim to test whether the no-arbitrage condition holds on the crude oil market. A naive approach would be to investigate the presence of cointegration between spot and future prices. But as argued by Brenner and Kroner (1995), they should not be cointegrated because commodity markets are subject to the so-called convenience yield, that is probably not a short memory process. Consequently the long-run equation is contaminated by an additive persistent component. However, as demonstrated by Liu and Tang (2010), the no-arbitrage condition can exist on commodity market if the convenience yield is non-negative and the no-arbitrage issue remains unsolved. Recently, Rossi and Santucci de Magistris (2013), have strengthened that the no-arbitrage condition between spot and future asset prices implies an analogous condition on their underlying volatilities. This original approach is convenient because it can be adapted to commodity markets under mild conditions on the volatility of the convenience yield. Accordingly, we aim to test for the no-arbitrage condition between the crude oil spot and futures volatilities rather than spot and future prices. The unbalanced cointegration framework is particularly appropriated here because we find some evidence that the persistent nature of the volatility can vary with the future maturity (i.e. unbalanced integration orders between spot and future volatility series).

The rest of the paper is laid out as follows. In Section 2, we introduce a bivariate stationary model of unbalanced cointegration. In Section 3, we develop the local Whittle estimator of unbalanced fractional cointegration (LWE-UFC). In Section 4 we demonstrate the consistency of the proposed estimator. Finite sample properties are investigated in Section 5. The application to the no-arbitrage condition between the volatilities of the spot and futures prices is proposed in Section 6. The Section 7 concludes the paper. Proofs are given in 8. Additional results and simulations are reported in 10 and 11.

[4] Hualde (2014) interestingly finds that the limiting distributions of estimates depend on a modified version of the Type II fractional Brownian motion. Different properties of this new type of fractional Brownian motion are discussed in Hualde (2012).

2. A stationary model of unbalanced cointegration

In the following, we say that a stochastic process ζ_t, has long memory $\alpha \in (0, 1/2)$, if its spectral density $f_\zeta(\lambda)$ satisfies $f_\zeta(\lambda) \sim g\lambda^{-2\alpha}$ as the frequency λ tends to 0, where the notation $\sim$ means that the ratio of the left and right sides tends to 1 in the limit. Then, ζ_t has short memory when $\alpha = 0$ and intermediate memory when $\alpha \in (-1/2, 0)$.[5] Now, we consider an unbalanced bivariate form of the triangular system introduced in Phillips (1991) and extended to the fractional framework by Nielsen (2004). Let y_t and x_t be two unbalanced observable variables with unknown real integration orders, δ_2 and $\delta_2 + \xi_n$ respectively. Hualde (2006) states that y_t and x_t are weakly unbalanced when δ_2 and $\delta_2 + \xi_n$ does not diverge at infinity (i.e. $\xi_n \to 0$ as $n \to \infty$) and strongly unbalanced when $\xi_n = |\xi| > 0$ as $n \to \infty$. To simplify notation, we shall use ξ to denote ξ_n. Then, the triangular unbalanced cointegration system is defined by

$$y_t = \beta x_t(\xi) + u_{1t}(-\delta_1), \qquad x_t = u_{2t}(-\delta_2 - \xi), \qquad t = 1, 2, ..., n, \tag{1}$$

where generically, $(-\alpha)$ denotes the fractional filter $(1 - L)^{-\alpha} = \sum_{k=0}^{\infty} a_k(\alpha)L^k$ with $a_k(\alpha) := \Gamma(k + \alpha)(\Gamma(\alpha)k!)^{-1}$, L, the lag operator and $\Gamma(.)$, the gamma function. Although standard cointegration theory does not apply to y_t and x_t , it does to y_t and $x_t(\xi)$. Thereby, the System (1) is a cointegration system in the sense that both series have a dominant common component with memory δ_2 that can be suitably recovered by filtering $x_t \sim I(\delta_2 + \xi)$ to obtain $x_t(\xi) \sim I(\delta_2)$.

Assumption 1. *y_t, x_t and $y_t - \beta x_t(\xi)$ are covariance stationary processes integrated of orders δ_2, $\delta_2 + \xi$ and δ_1 respectively, and satisfying*

$$0 \le \delta_1 < \delta_2 < \delta_2 + |\xi| < 1/2 \tag{2}$$

where $|\xi| < k$, with k an arbitrary real number small compared to δ_2.

Under Assumption 1, the System in (1) provides a valid data-generating process for stationary cointegration model so that $z_t = (y_t - \beta x_t(\xi), x_t)'$ possesses a spectral density, $f_z(\lambda_j)$, where λ_j denotes the Fourier frequencies, $\lambda_j = 2\pi j/n$, with $j = 1, \ldots, m$ and $m = o(n)$, the bandwidth parameter. Assumption 1 leaves out anti-persistent processes because they clearly have limited economic relevance.

Now, assume that $u_t = (u_{1t}, u_{2t})'$ has short memory with spectral density $f_u(\lambda_j)$ satisfying, $f_u(\lambda_j) \sim G$ in the neighborhood of the origin, with G (the long run covariance matrix) a real, symmetric, finite and positive definite matrix. Notice that when cointegration arises, rank$(G) < 2$, so that G has reduced-rank whether or not $\xi \neq 0$ (see Hualde 2006).

[5]When $\alpha \in [1/2, 1]$, the spectral density of ζ_t is no longer defined although Velasco (1999b) demonstrated it has a pseudo-spectral density and standard local Whittle-based estimators are biased. In such a case, Velasco (1999b) proposed to use a tapered periodogram to reduce the bias, but at the cost of a higher variance. Exploiting this result, Shimotsu (2012) combined a multivariate extension of the exact local Whittle estimator of Shimotsu (2010) and a tapered version of the Robinson (2008) to propose a fractional cointegration estimator which is consistent over $-1/2 < \delta_1 < \delta_2 < \infty$ and includes $\delta_2 = 1$ and $\delta_1 = 0$ as a special case.

Remark 1. *Assumptions on u_t entails only mild conditions so that u_t is possibly a vector ARMA process or any other short memory processes with a Wold representation, $u_t = C(L)\varepsilon_t$, where ε_t are further defined as martingale difference innovations and $C(L)$ is a square-summable causal matrix filter satisfying $G = C(1)C(1)'(2\pi)^{-1}$.*

In the present study, we define processes only on the vicinity of the origin, in view of allowing a semi-parametric treatment of the short-run dynamics. Indeed, we support that such approach is particularly of interest in cointegration analysis because empirical interest is more likely to lie in long run dynamics. Accordingly, by Assumption 1 and considering the Remark 1, z_t has a spectral density, so that

$$E\Big(z_t - E(z_t)\Big)\left(z'_{t+k} - E(z'_t)\right) = \int_{-\pi}^{\pi} e^{ik\lambda} f_z(\lambda) d\lambda, \tag{3}$$

with $f_z(\lambda) = \Lambda(\lambda)^{-1} f_u(\lambda) \left(\Lambda(\lambda)^*\right)^{-1}$ and $\Lambda(\lambda) = \text{diag}\left((1-e^{i\lambda})^{\delta_1}, (1-e^{i\lambda})^{\delta_2+\xi}\right)$. Since $(1-e^{i\lambda})^{\alpha} = (|2\sin(\lambda/2)|)^{\alpha} \sim \lambda^{\alpha}$ as $\lambda \to 0$, we can avoid a parametric treatment of $f_z(\lambda)$ in favor of the following local power law representation around zero frequency,

$$f_z(\lambda) \sim \Big(\Lambda(\lambda;\vartheta_1)\Big)^{-1} G \Big(\Lambda(\lambda;\vartheta_1)^*\Big)^{-1}, \qquad \Lambda(\lambda;\vartheta_1) = \text{diag}\left(\lambda^{\delta_1}, \lambda^{\delta_2+\xi}\right), \qquad \text{as } \lambda \to 0 \tag{4}$$

where $\vartheta_1 = (\delta_1, \delta_2 + \xi)'$ and the superscript $*$ denotes the conjugate transpose. As Nielsen (2007) we assume that G is diagonal so that u_{1t} and u_{2t} are incoherent in the vicinity of the origin and the Equation (4) is correctly specified.[6] Thereby, the phase parameter modeled as $\varphi = (\pi - \lambda)(\delta_2 - \delta_1)/2$ in Robinson (2008) and Shimotsu (2012) is null in our framework (see also Shimotsu 2007).[7] Notice also that, if $\delta_2 \leq \delta_1$, β cannot be identified. This issue also occurs in standard regression analysis of balanced cointegration.

A matrix representation of the System (1) gives

$$\begin{pmatrix} 1 & -\beta(1-L)^{\xi} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} (1-L)^{-\delta_1} & 0 \\ 0 & (1-L)^{-\delta_2-\xi} \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}. \tag{5}$$

The local Whittle analysis of restricted versions of Equation (5) is not new. For instance, when $\beta = 0$, y_t and x_t do not share any long-run component, and the first matrix on the left side of (5) reduces to the identity matrix. Thereby, one can recover the so-called stationary ARFIMA model whose the local Whittle estimate has been studied in univariate framework by Robinson (1995a) and later extended to a multivariate setting by Lobato (1999) and Shimotsu (2007). When $\beta \neq 0$ and $\xi = 0$, cointegration can arise in the usual sense. Considering this latter case, Robinson and Marinucci (2003) and Nielsen (2005) discuss the estimation of β by means of frequency domain least squares (see also Robinson 1994, Christensen and Nielsen 2006, Nielsen and Frederiksen 2011). As mention previously, approximate local

[6]Robinson (2008) propose a local Whittle treatment of multivariate stationary systems with unknown phase, considering the model of Nielsen (2007) as a special case.

[7]The presence of non-null off-diagonal elements in G should imply non-negligible imaginary part of the cross-spectrum element $f_z^{ab}(\lambda)$ such as $f_z^{ab}(\lambda) \sim G_{ab}\lambda^{-\delta_a-\delta_b}e^{i(\pi-\lambda)(\delta_a-\delta_b)/2}$ as $\lambda \to 0$, for $a, b = 1, 2$ and where G_{ab} denotes the (a, b)th element of G.

Whittle estimation of δ_1, δ_2 and β is considered in Nielsen (2007), Robinson (2008) and Shimotsu (2012). When $\xi \neq 0$ and $\delta_2 > 1/2$, Hualde (2014) derives the asymptotic properties of the non-linear least squares estimator of β and ξ.

3. Local Whittle estimation

In the following, we introduce a local Whittle estimator of $\theta = (\delta_1, \delta_2, \xi, \beta)'$. Let, I_z be the periodogram matrix of z_t, defined as $I_z(\lambda_j; \vartheta_2) = w_z(\lambda_j; \vartheta_2) w_z(\lambda_j; \vartheta_2)^*$ with $w_z(\lambda_j; \vartheta_2) = (2\pi n)^{-1/2} \sum_{t=1}^{n} z_t e^{it\lambda_j}$, the Fourier transform of z_t and $\vartheta_2 = (\beta, \xi)'$. Updated to bandwidth $m = o(n)$, i.e. $j = 1, \ldots, m$, in view of the local treatment we obtain,

$$I_z(\lambda_j; \vartheta_2) = \begin{pmatrix} w_y(\lambda_j) - \beta \lambda_j^{\xi} w_x(\lambda_j) \\ w_x(\lambda_j) \end{pmatrix} \begin{pmatrix} w_y(\lambda_j) - \beta \lambda_j^{\xi} w_x(\lambda_j) \\ w_x(\lambda_j) \end{pmatrix}^* . \tag{6}$$

Thereby, the presence of λ_j^{ξ} corrects for the fact that long memory parameters of y_t and x_t are unbalanced. Then, the discrete local Whittle approximation to the likelihood is

$$\begin{aligned} Q_m(\theta, G) = m^{-1} \sum_{j=1}^{m} \Big[&\log\det \left(\left(\Lambda(\lambda_j; \vartheta_1) \right)^{-1} G \left(\Lambda(\lambda_j; \vartheta_1)^* \right)^{-1} \right) \\ &+ \operatorname{tr} \left(G^{-1} \Lambda(\lambda_j; \vartheta_1) I_z(\lambda_j; \vartheta_2) \Lambda(\lambda_j; \vartheta_1)^* \right) \Big] , \end{aligned} \tag{7}$$

where $G \in \Theta_G$, the set of real positive definite 2×2 matrices. The objective function Q is minimized over Θ_G by

$$\hat{G}(\theta) = \operatorname{Re} \left(m^{-1} \sum_{j=1}^{m} \Lambda(\lambda_j; \vartheta_1) I_z(\lambda_j; \vartheta_2) \Lambda(\lambda_j; \vartheta_1)^* \right), \tag{8}$$

leading to the following concentrated likelihood function

$$R_m(\theta) = \log\det \hat{G}(\theta) - \frac{2(\delta_1 + \delta_2 + \xi)}{m} \sum_{j=1}^{m} \log \lambda_j. \tag{9}$$

Accordingly, the local Whittle estimator of θ is defined as $\hat{\theta} = \arg\min_{\theta \in \Theta} R_m(\theta)$, for $m \in [1, n/2]$ and Θ a compact subset of $\mathbb{R}^4$ with $\Theta = \Theta_\delta \times \Theta_\xi \times \Theta_\beta$ and $\delta = (\delta_1, \delta_2)'$. Vectors θ and $\hat{\theta}$ are respectively the vector of unknown and estimated values, $(\delta_1, \delta_2, \xi, \beta)'$ and $(\hat{\delta}_1, \hat{\delta}_2, \hat{\xi}, \hat{\beta})'$. Observe that Equation (8) yields $\hat{G}_{11}(\theta) = \operatorname{Re} \left(m^{-1} \sum_{j=1}^{m} \lambda_j^{2\delta_1} I_z^{11}(\lambda_j; \vartheta_2) \right)$ with $I_z^{11}(\lambda_j; \vartheta_2) = I_{yy}(\lambda_j) - 2\beta \lambda_j^{\xi} I_{xy}(\lambda_j) + \beta^2 \lambda_j^{2\xi} I_{xx}(\lambda_j)$, which has some similarities with the weighted least squares of Nielsen (2005). In 10, we show that the Proposition 1 of Nielsen (2005, p. 297) and the Theorem 1 of Robinson and Marinucci (2003) remain valid when $\xi \neq 0$ if x_t is appropriately differenced. Accordingly, we anticipate that $\hat{\beta}$ is also $\lambda_m^{\delta_{01} - \delta_{02}}$-consistent when $\xi \neq 0$.

4. Consistency

To prove the consistency of this local Whittle estimator, we introduce several assumptions, fairly similar to those of Shimotsu (2007) and Nielsen (2007). In the following, θ_0 and G_0 will denote the true parameter values of θ and G. Then, let $f_z^{ab}(\lambda)$ and G_{ab}^0 denote the (a,b)th element of $f_z(\lambda)$ and G_0 respectively. Define also $\vartheta_{01} = (\delta_{01}, \delta_{02} + \xi_0)'$ and δ_{0a} the ath element of ϑ_{01}.

Assumption 2. *As $\lambda \to 0^+$, elements of the spectral density $f_z(\lambda)$ satisfies*

$$f_z^{ab}(\lambda) = G_{ab}^0 \lambda^{-\delta_{0a}-\delta_{0b}} + o(\lambda^{-\delta_{0a}-\delta_{0b}}), \qquad a,b = \{1,2\}, \tag{10}$$

where matrix G is finite, real, symmetric and positive definite. Also assume $G_{12} = G_{21} = 0$.

Assumption 3. *The sequence z_t is a linear process defined as*

$$z_t - E(z_t) = A(L)\varepsilon_t = \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}, \qquad \sum_{j=0}^{\infty} ||A_j||^2 < \infty, \tag{11}$$

with $||.||$ the Euclidean norm, so that A_j is a causal square summable matrix filter. Moreover, ε_t satisfies, almost surely, $E(\varepsilon_t|F_{t-1}) = 0$, $E(\varepsilon_t \varepsilon_t'|F_{t-1}) = I_2$, with F_t a σ-field generated by $\{\varepsilon_s, s \leq t\}$ and there exist a random variable ε such that $E(\varepsilon^2 < \infty)$ and for all $\eta > 0$ and some $K > 0$, $Pr(||\varepsilon_t||^2 > \eta) \leq KPr(\varepsilon^2 > \eta)$.

Assumption 2 implies a zero-coherence condition that applies only in the vicinity of the origin. As argued in Nielsen (2007), it is a less restrictive assumption than the traditional orthogonality condition encountered in the least squares theory. Notably, it allows for errors to be correlated away from the origin and share, for instance, a common short- and/or medium-term dynamics. As mentioned previously, Robinson (2008) and Shimotsu (2012) relax this hypothesis. The present estimator could be modified to model f_z correctly when G^{12} is non-null by specifying the phase parameter and presumably, its consistency could be demonstrated in this case but we do not pursue that possibility further. Assumption 3 imposes uniformly square integrable martingale-difference innovations with constant conditional variance in view of the application of the standard CLT for martingale-difference arrays.[8] The latter assumption on conditional variance could also probably be relaxed assuming boundedness of higher moments as in Robinson and Henry (1999) but we do not investigate this issue further.

Assumption 4. *In a neighborhood of the origin, $A(\lambda) = \sum_{j=0}^{\infty} A_j e^{ij\lambda}$ is differentiable and*

$$\frac{\partial}{\partial \lambda} A_a(\lambda) = O(\lambda^{-1}||A_a(\lambda)||) \text{ as } \lambda \to 0^+ \tag{12}$$

where $A_a(\lambda)$ is the a-th row of $A(\lambda)$.

[8]We aim to investigate the asymptotic distribution of the estimator but for now the paper only discusses the consistency.

Assumption 4 implies $\partial A_a(\lambda)/\partial\lambda = O(\lambda^{-\delta_a-1})$ because by the Cauchy inequality

$$||A_a(\lambda)|| \leq (A_a(\lambda)A_a^*(\lambda))^{1/2} = (2\pi f_{aa}(\lambda))^{1/2}.$$

Thereby, under Assumption 3 and 4 we have $f_z(\lambda) = (2\pi)^{-1}A(\lambda)A(\lambda)^*$.

Assumption 5. *As $n \to \infty$, the bandwidth parameter satisfies*

$$\frac{1}{m} + \frac{m}{n} \to 0. \tag{13}$$

where $m = \lfloor n^k \rfloor$, $k \in (0, 4/5]$.

The bandwidth requirement defined by the Assumption 5 ensures that m tends to ∞ as $n \to \infty$ but at a slower rate to remain in a neighborhood of the origin. The bandwidth parameter m is theoretically bounded by $n^{4/5}$ but in practice a too small bandwidth increases the variance of the estimator while a too large m generally increases the bias. Accordingly, Assumptions 2-5 are analogous to Assumptions 1-4 of Shimotsu (2007), Nielsen (2007) and natural multivariate extensions of Assumptions A1-A4 of Robinson (1995a). Under Assumptions 2-5 we may now state the following theorem which establishes the convergence rate of $\hat{\theta}$.

Theorem 1. *Let Assumption 1-5 hold. Define $\nu_0 = \delta_{02} - \delta_{01}$. Then, for $\theta_0 \in \Theta$ and $\delta_{01} < \delta_{02} \leq \delta_{02} + |\xi_0|$, as $n \to \infty$,*

$$\begin{pmatrix} \hat{\delta}_1 - \delta_{01} \\ \hat{\delta}_2 - \delta_{02} \end{pmatrix} = O_p(m^{-1/2}) \tag{14}$$

$$(\hat{\xi} - \xi_0) = O_p(m^{-1/2}) \tag{15}$$

$$(\hat{\beta} - \beta_0) = O_p(m^{-1/2}(n/m)^{-\nu_0}) \tag{16}$$

The convergence rate of $\hat{\beta}$ confirms that the system is rebalanced when x_t is appropriately filtered. Notice that when $\delta_1 \to 0$ and $\delta_2 \to 1/2$, $\hat{\beta}$ is almost $\sqrt{n}$-consistent because $m^{1/2}m^{\delta_1-\delta_2\to-1/2}n^{\delta_2-\delta_1\to1/2}(\hat{\beta} - \beta_0) \xrightarrow{p} 0$. More importantly, when the presence of $\xi \neq 0$ is neglected and thus $x_t(\xi)$ is trivially replaced by x_t in Equation (1), $\hat{\beta} - \beta_0 = O_p((n/m)^{\delta_1-(\delta_2+\xi)})$ and $\hat{\beta}$ is likely to be inconsistent if $\xi < 0$.

5. Monte Carlo experiment

5.1. Simulation design

This section discusses the finite sample performance of the proposed estimator by means of Monte Carlo simulations. As argued in Hurvich and Ray (1995), it is not so easy for practitioners to identify

Table 1: Simulation results for the stationary model when $\xi = 0.1$ and $\rho = 0$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.4	0	δ_2	0.020	0.031	0.176	0.034	0.019	0.142	0.034	0.012	0.115
		δ_1	0.063	0.039	0.207	0.074	0.025	0.175	0.076	0.017	0.150
		ξ	-0.014	0.005	0.070	-0.015	0.002	0.042	-0.012	0.001	0.027
		β	0.016	0.006	0.079	0.022	0.004	0.068	0.029	0.003	0.060
0.4	0.2	δ_2	-0.013	0.035	0.188	-0.003	0.022	0.150	-0.002	0.015	0.122
		δ_1	0.008	0.031	0.176	0.009	0.019	0.139	0.014	0.011	0.105
		ξ	-0.003	0.012	0.109	-0.010	0.006	0.076	-0.013	0.003	0.052
		β	0.006	0.019	0.139	0.025	0.012	0.113	0.041	0.007	0.092
0.4	0.3	δ_2	-0.029	0.044	0.212	-0.021	0.028	0.170	-0.014	0.018	0.135
		δ_1	-0.006	0.030	0.175	-0.005	0.019	0.138	-0.005	0.011	0.104
		ξ	0.009	0.020	0.143	0.003	0.011	0.106	-0.004	0.006	0.079
		β	-0.025	0.035	0.189	0.002	0.022	0.149	0.024	0.013	0.118
$m = \lfloor n^{0.8} \rfloor$											
0.4	0	δ_2	-0.041	0.004	0.072	-0.044	0.002	0.061	-0.052	0.001	0.062
		δ_1	0.033	0.004	0.072	0.054	0.002	0.073	0.076	0.002	0.086
		ξ	0.016	0.002	0.043	0.014	0.001	0.029	0.015	0.000	0.021
		β	0.057	0.001	0.068	0.071	0.001	0.078	0.086	0.001	0.090
0.4	0.2	δ_2	-0.044	0.005	0.081	-0.041	0.002	0.063	-0.041	0.001	0.056
		δ_1	-0.006	0.003	0.059	0.000	0.002	0.042	0.005	0.001	0.032
		ξ	0.022	0.003	0.057	0.021	0.001	0.041	0.020	0.001	0.032
		β	0.065	0.002	0.081	0.074	0.001	0.083	0.081	0.001	0.087
0.4	0.3	δ_2	-0.034	0.006	0.084	-0.029	0.003	0.062	-0.025	0.002	0.048
		δ_1	-0.016	0.004	0.064	-0.011	0.002	0.045	-0.007	0.001	0.034
		ξ	0.015	0.004	0.064	0.015	0.002	0.046	0.014	0.001	0.035
		β	0.049	0.004	0.079	0.055	0.002	0.073	0.059	0.002	0.071

whether or not the observed time series are stationary. We deal with this issue by considering a data generating process (DGP) that accommodates the mean-reverting non-stationary regions of the parameter space, although the developed estimator is not theoretically designed to handle this case. Thereby, we generate a fractionally cointegrated system according to the following model,

$$y_t = \beta x_t(\xi) + u^{\#}_{1t}(-\delta_1), \qquad x_t = u^{\#}_{2t}(-\delta_2 - \xi), \qquad t = 1, 2, ..., n, \tag{17}$$

where for a generic process ζ_t, $\zeta^{\#}_t = \zeta_t I(t \geq 1)$, with $I(.)$ the indicator function. $\zeta^{\#}_t(-\alpha)$ denotes the fractional truncated filter $\zeta^{\#}_t(-\alpha) := \sum_{k=0}^{t-1} \Gamma(k+\alpha)(\Gamma(\alpha)k!)^{-1}\zeta_{t-k}$. The System in (17) provides a valid DGP for both, stationary and non-stationary regions of the parameter space, given that y_t and x_t are type II processes, thereby contrasting with our system in (1) which is of type I. Marinucci and Robinson (1999) shown that type I and type II processes are asymptotically equivalent in the stationary regions of the parameter space if both are generated from the same short memory sequence (see also Robinson 2005). Accordingly, the type II representation is often retained in simulation study for its simplicity. Nonetheless, Davidson and Hashimzade (2009) demonstrated that both representations can substantially differ in finite sample. We account for their recommendations by performing some additional simulations

Table 2: Simulation results for the stationary model when $\xi = 0.1$ and $\rho = 0.4$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.4	0	δ_2	0.038	0.026	0.167	0.045	0.015	0.131	0.038	0.009	0.101
		δ_1	0.035	0.033	0.184	0.034	0.020	0.144	0.039	0.011	0.112
		ξ	0.029	0.004	0.067	0.030	0.001	0.046	0.031	0.001	0.039
		β	0.315	0.009	0.329	0.323	0.006	0.332	0.331	0.004	0.336
0.4	0.2	δ_2	-0.003	0.037	0.192	0.011	0.023	0.152	0.015	0.014	0.118
		δ_1	-0.006	0.032	0.180	-0.013	0.020	0.141	-0.016	0.012	0.109
		ξ	0.021	0.010	0.100	0.012	0.004	0.062	0.008	0.002	0.041
		β	0.366	0.023	0.396	0.383	0.014	0.401	0.396	0.009	0.406
0.4	0.3	δ_2	-0.024	0.043	0.209	-0.015	0.029	0.171	0.001	0.018	0.135
		δ_1	-0.008	0.033	0.181	-0.016	0.020	0.144	-0.018	0.012	0.110
		ξ	0.020	0.015	0.124	0.012	0.009	0.094	0.001	0.004	0.064
		β	0.357	0.046	0.416	0.387	0.026	0.420	0.403	0.016	0.423
$m = \lfloor n^{0.8} \rfloor$											
0.4	0	δ_2	-0.072	0.003	0.092	-0.066	0.002	0.077	-0.067	0.001	0.074
		δ_1	0.037	0.003	0.069	0.043	0.002	0.062	0.054	0.001	0.066
		ξ	0.079	0.002	0.089	0.073	0.001	0.077	0.064	0.000	0.066
		β	0.378	0.002	0.380	0.397	0.001	0.399	0.415	0.001	0.416
0.4	0.2	δ_2	-0.061	0.003	0.083	-0.054	0.002	0.069	-0.053	0.001	0.062
		δ_1	0.001	0.003	0.054	0.002	0.002	0.039	0.002	0.001	0.029
		ξ	0.054	0.002	0.071	0.050	0.001	0.059	0.046	0.000	0.050
		β	0.424	0.003	0.427	0.432	0.002	0.434	0.440	0.001	0.441
0.4	0.3	δ_2	-0.042	0.004	0.076	-0.037	0.002	0.058	-0.032	0.001	0.047
		δ_1	-0.010	0.003	0.057	-0.007	0.002	0.041	-0.006	0.001	0.031
		ξ	0.030	0.002	0.057	0.029	0.001	0.045	0.027	0.001	0.037
		β	0.427	0.004	0.432	0.433	0.002	0.436	0.436	0.002	0.438

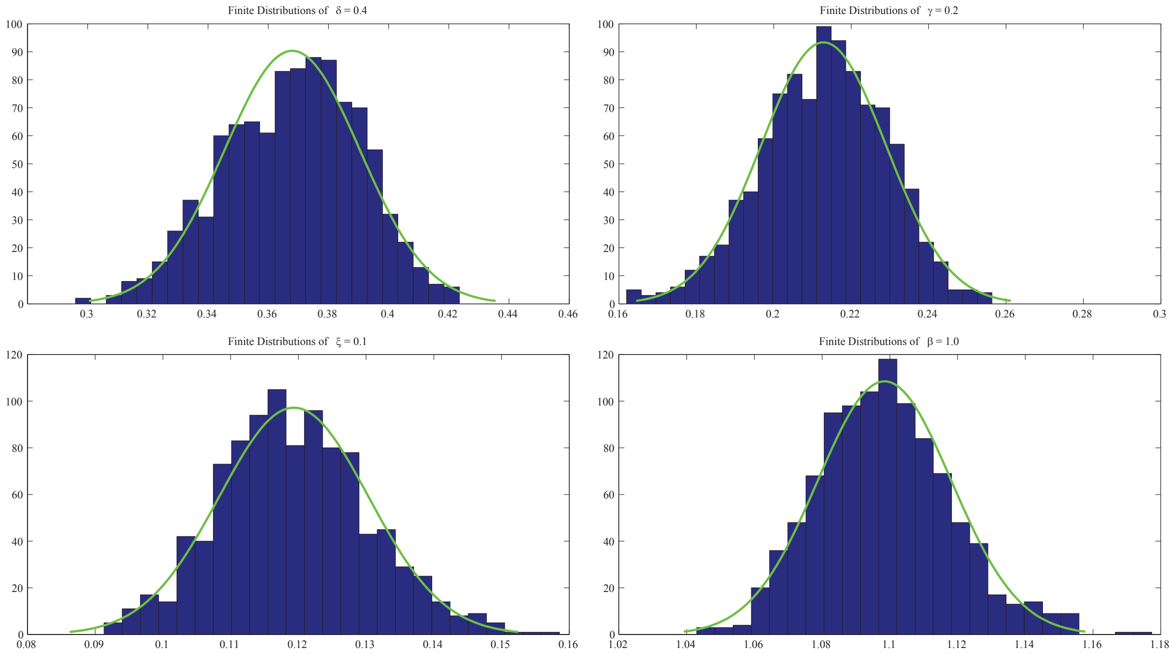
with a type I representation. Davidson and Hashimzade (2009) discuss the benefits and limitations of several techniques devoted to type I simulation. Given that we are interested in generating simple linear fractional processes, we use the circulant embedding method extended to multivariate fractional Gaussian noise by Helgason et al. (2011).

In our experiment we consider the following settings. The stationary cointegration case is explored for $\delta_2 = 0.4$ and $\delta_1 = \{0.0, 0.2, 0.3\}$. Similarly, the strong and weak cointegration cases, are investigated for $\delta_1 = \{0.0, 0.2, 0.4\}$ and $\delta_1 = \{0.4, 0.6, 0.8\}$ respectively, with $\delta_2 = \{0.6, 0.8, 1.0\}$.[9] Because in practice the weakly unbalanced cointegration case is generally indistinguishable from that of balanced cointegration, we do not consider this case in the simulation. Conversely, the strongly unbalanced cointegration case has greater applicability and is investigated for $\xi = \{-0.1, 0.1\}$. The long-run coefficient is fixed as $\beta = 1$. The vector u_t is generated from a bivariate normal distribution $N_2(\mu, \Sigma)$ distribution. For each simulation, we report the bias, the variance and the Root Mean Squared Error (RMSE), defined by $\frac{1}{I}\sum_{i=1}^{I} E\left((\hat{\theta}_i - \theta)^2\right) := Var(\hat{\theta}) + Bias(\hat{\theta}|\theta)^2$ with I the number of replications set to 10000. Following Shimotsu (2012), we use a penalty term defined as $\Pi(\beta, \tilde{\beta}) = \min(0, \beta - \tilde{\beta} + C)^4 + \max(0, \beta - \tilde{\beta} - C)^4$ to govern the objective function

[9]To generate nonstationary series (e.g. x_t with $\delta_2 \geq 1/2$), we simulate an intermediate stationary (or asymptotically stationary) process, integrated of order $I(\delta_\Delta = \delta_2 - 1)$ and cumulate the resulting series.

when the space for dimensionality reduction is weak, with $\tilde{\beta} = \tilde{\beta}_{LSE}$, so that it preserves the asymptotic results obtained in Equation (2). Alternatively, the Narrow-Band Least Squares (NBLS) estimate has also been used for initialization and does not significantly modified the results. The initial estimates for $\tilde{\delta}_x$ and $\tilde{\delta}_y$ are obtained from the local Whittle estimator of Robinson (1995a) applied to x_t and y_t respectively. Therefore, the initial estimate for ξ, namely $\tilde{\xi}$, is based on the difference between $\tilde{\delta}_y$ and $\tilde{\delta}_x$ (see Hualde 2006, p. 777). The initial estimate, $\tilde{\beta}_{LSE}$, results from the regression of y_t and $x_t(\tilde{\xi})$. All computations are performed using MATLAB 2013a.

Figure 1: Finite sample distribution of $\hat{\theta}$ for 1000 replications of the stationary cointegration model with $\rho = 0$ and $n = 4096$.



5.2. *Simulation results*

Table 1 reports the results for 10000 replications of the stationary cointegration model when $\xi = 0.1$ and $\rho = 0$. Accordingly, this specification satisfies all the assumptions of the model defined in Equation (1). In absence of short-run dynamics, taking frequencies away from the origin essentially impact the variance (rather than the bias) compare with $m = \lfloor n^{0.5} \rfloor$. Thereby, in both cases, the approximation in Equation (4) is close to $f_z(\lambda) = (2\pi)^{-1}A(\lambda)A(\lambda)^*$. Observe that the strength of the cointegration and the variance of $\hat{\beta}$ are negatively linked which is in accordance with the limit theory (the convergence rate of $\hat{\beta}$ depends on the strength of the cointegration). In all cases, the variances (the RMSE so on) decrease as the sample size increases. Moreover, variances are lower when $m = \lfloor n^{0.8} \rfloor$.

Table 2 reports the results for 10000 replications of the stationary cointegration model when $\xi = 0.1$ and $\rho = 0.4$. Thereby, the off-diagonal elements of G are non-null and a correlation between $y_t - \beta x_t(\xi)$

Table 3: Comparison of type I and II processes when $\xi = -0.1$, $\rho = 0$ and $m = \lfloor n^{0.8} \rfloor$

Type I			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.4	0	δ_2	-0.059	0.029	0.180	-0.011	0.017	0.130	0.021	0.005	0.075
		δ_1	0.045	0.017	0.138	0.050	0.013	0.126	0.046	0.008	0.100
		ξ	0.136	0.015	0.183	0.101	0.009	0.137	0.078	0.003	0.093
		β	-0.094	0.002	0.106	-0.035	0.002	0.055	0.046	0.001	0.056
0.4	0.2	δ_2	-0.108	0.008	0.141	-0.088	0.004	0.109	-0.074	0.002	0.088
		δ_1	-0.041	0.004	0.076	-0.037	0.002	0.061	-0.030	0.001	0.047
		ξ	0.126	0.003	0.136	0.114	0.001	0.119	0.105	0.001	0.108
		β	0.006	0.004	0.060	0.044	0.002	0.061	0.068	0.001	0.075
0.4	0.3	δ_2	-0.020	0.006	0.081	-0.018	0.003	0.055	-0.013	0.002	0.042
		δ_1	-0.020	0.004	0.066	-0.013	0.002	0.047	-0.010	0.001	0.034
		ξ	0.006	0.002	0.049	0.007	0.001	0.033	0.006	0.001	0.024
		β	0.043	0.004	0.076	0.052	0.002	0.072	0.056	0.002	0.068
Type II											
0.4	0	δ_2	-0.033	0.004	0.069	-0.036	0.002	0.058	-0.039	0.001	0.053
		δ_1	0.029	0.004	0.071	0.043	0.002	0.065	0.055	0.002	0.068
		ξ	0.017	0.001	0.042	0.021	0.001	0.033	0.024	0.000	0.028
		β	0.050	0.001	0.063	0.067	0.001	0.074	0.083	0.001	0.087
0.4	0.2	δ_2	-0.035	0.005	0.078	-0.035	0.003	0.062	-0.037	0.002	0.054
		δ_1	-0.002	0.004	0.064	0.006	0.002	0.046	0.011	0.001	0.036
		ξ	0.016	0.003	0.054	0.018	0.001	0.041	0.020	0.001	0.033
		β	0.060	0.002	0.077	0.069	0.001	0.079	0.078	0.001	0.085
0.4	0.3	δ_2	-0.030	0.006	0.083	-0.029	0.003	0.063	-0.029	0.002	0.050
		δ_1	-0.010	0.004	0.065	-0.004	0.002	0.047	0.000	0.001	0.034
		ξ	0.010	0.004	0.063	0.013	0.002	0.046	0.014	0.001	0.035
		β	0.043	0.004	0.076	0.051	0.002	0.070	0.056	0.002	0.069

and the regressor is introduced at all frequencies. Consequently, Assumption 2 is now violated and the estimator faces an additional complication. Nonetheless, simulations do not explicitly reproduce these theoretical results. For instance, the long memory parameters are estimated fairly precisely, especially when $m = \lfloor n^{0.8} \rfloor$.

Concerning $\xi < 0$, we only report the stationary cointegration case with $\rho = 0$ because the results are very similar as we can see in the lower part of the Table 3. The simulations based on the type I representation are reported in the upper part of the Table 3. We limit our investigation to this case because the computations are highly time consuming but we support that these results are sufficiently informative. We can observe that the differences with the type II representation are sometimes substantial. For instance, whatever the sample, the bias slightly differs when $\delta_1 = 0$. In some cases, the variance increases but the divergences disappear when $\delta_1 = 3$.

Tables 7, 8, 9 and 10 report the results for strong and weak fractional cointegration cases with either $\rho = 0$ or $\rho \neq 0$. In all cases, there is weak evidence of consistency given that RMSEs slowly decrease when the sample size increases. This result suggests that practitioners should devote a particular attention to the stationarity or non-stationarity of the data. In the latter case, they should use the estimator of Hualde (2014).

6. Empirical illustration

In this section we aim to test whether the no-arbitrage condition holds on the crude oil market. A naive approach would be to investigate the presence of cointegration between spot and future prices. But as argued by Brenner and Kroner (1995), they should not be cointegrated because commodity markets are subject to the so-called convenience yield, that is probably not a short memory process. Consequently the long-run equation is contaminated by an additive persistent component. However, as demonstrated by Liu and Tang (2010), the no-arbitrage condition can exist on commodity market if the convenience yield is non-negative and the no-arbitrage issue remains unsolved. Recently, Rossi and Santucci de Magistris (2013), have strengthened that the no-arbitrage condition between spot and future asset prices implies an analogous condition on their underlying volatilities. This original approach is convenient because it can be adapted to commodity markets under mild conditions on the volatility of the convenience yield. Accordingly, we aim to test for the no-arbitrage condition between the crude oil spot and futures volatilities rather than spot and future prices.

The no-arbitrage condition states that on financial markets, risk-free arbitrage opportunities cannot arise. Paradoxically, Grossman and Stiglitz (1980) emphasized that arbitrage opportunities, even infrequent, are necessary to make the market sufficiently incentive for market participants. Investigating the no-arbitrage condition is also interesting because it is closely related to the efficient market hypothesis issue. Denoting $F_{t+k|t}$ and S_t the futures and the spot prices respectively, the no-arbitrage condition implies $F_{t+k|t} = S_t \cdot e^{k \cdot r_{t+k|t}}$ with $r_{t+k|t}$ the return of a risk-free asset that expires at time $t+k$ and $e^{k \cdot r_{t+k|t}}$ the cost of carry premium. Rossi and Santucci de Magistris (2013) show that using the rescaled daily range, $\sigma_{t,X} = (\log 2)^{-1/2}(\max_\tau \log X_\tau - \min_\tau \log X_\tau)$, for X any price, the no-arbitrage condition is directly related with the second moments of the spot and futures prices by the following equation,

$$\sigma_{t,F} = \sigma_{t,S} + b_t + u_t, \qquad b_t = (\log 2)^{-1/2}(r_{\tau_{\max}} - r_{\tau_{\min}}), \qquad t-1 < \tau \leq t \tag{18}$$

where $r_{\tau_{\max}}$ and $r_{\tau_{\min}}$ stand for the risk-free rate in correspondence, respectively, of the highest and lowest log-price in a given day. The additional term u_t is justified by the presence of market frictions whereas b_t could be omitted because the risk-free asset intraday variations are negligible.

Unfortunately, Equation (18) is not correct with respect to our framework because we focus on commodity markets whereas Rossi and Santucci de Magistris (2013) focus on stock index futures and thus neglect the so-called convenience yield. The convenience yield is defined as “the flow of benefit of immediate ownership of a physical commodity” (see Liu and Tang 2011). Reformulating the no-arbitrage condition we obtain

$$F_{t+k|t} = S_t \cdot e^{k \cdot (r_{t+k|t} - c_{t+k|t})} \tag{19}$$

where $c_{t+k|t}$ represents the net convenience yield that is the difference between the convenience yield and the cost of storage applicable from time t to time $t+k$. Following the same reasoning as Rossi and Santucci de Magistris (2013), our modified version of Equation (18) should hold if the intraday volatility of the convenience yield is negligible. However, the convenience yield is unobservable and this assumption requires some explanations. Implications of this hypothesis are twofold: (i) it ensures that Equation (18) is not contaminated by the presence of an additional volatility component; (ii) it reduces the probability for the convenience yield to be negative. Indeed, in the latter case, the futures prices would be too high relative to the spot price, thereby offering an arbitrage opportunity. In financial literature, the net convenience yield is often modeled as an Ornstrein-Uhlenbeck process. In such a representation the convenience yield volatility is generally high and negative values may frequently occur. Nonetheless, in a recent paper, Liu and Tang (2010) consider Cox-Ingersoll-Ross representation of the convenience yield and demonstrate that its volatility is no necessarily high. Accordingly, in line with their results, we will assume that the intraday volatility of the convenience yield is sufficiently low to be neglected in Equation (18) and to avoid frequent arbitrage-free opportunities.

In the following, we conjecture that when substituting the daily range by the daily squared returns, the relation in Equation (18) remains valid.[10] Accordingly, we can test whether the no-arbitrage condition holds, by estimating the Equation (18). Nonetheless, the persistent nature of the volatility reveals that long run components drive the underlying processes. In such a case, testing for the presence of cointegration is useful to guard against the risk of spurious regression but also because the dynamics of $\sigma_{t,F}$ and $\sigma_{t,S}$ are likely to slightly diverge in short run. Furthermore, one would expect that the more the maturity of the futures is far away, the more the volatilities are likely to drift far apart. Some evidences of this mechanism are provided by Caporale et al. (2014) on the spot and futures prices.

We focus on WTI crude oil spot and futures prices traded in NYMEX. Our data set runs from January 2, 1996 to December 16, 2013 for a total of $n = 4499$ observations.[11] We consider four different maturity contracts. The contract *Futures 1* specifies the earliest delivery date. It expires on the third business day prior to the 25th calendar day of the month preceding the delivery month. If the 25th calendar day of the month is a non-business day, trading ceases on the third business day prior to the business day preceding the 25th calendar day. The contracts *Futures 2-4* represent the successive delivery months following the contract *Futures 1*. Graphics 2 and 3 represents the daily log squared returns of the spot prices and the four futures.

To test for the presence of stationary (unbalanced) cointegration we apply a rigorous methodology. First of all, we estimate the long memory parameters of each volatility series. Because their is no particular reasons for the volatility to be stationary, we use the two-step exact local Whittle (2S-ELW) estimator of Shimotsu (2010) which is robust to the presence of unknown mean and polynomial time trend. The results are reported in Table 4 for several bandwidths. Clearly, they are fairly homogeneous for a given

[10]We employ this rough approximation because we do not have the data but we planned to solve this issue before submitting the paper.

[11]All data were collected on the website of the US Energy Information Administration: http://www.eia.gov/petroleum/.

Figure 2: Daily squared return volatility proxy of the crude oil spot prices from January 2, 1996 to December 16, 2013.

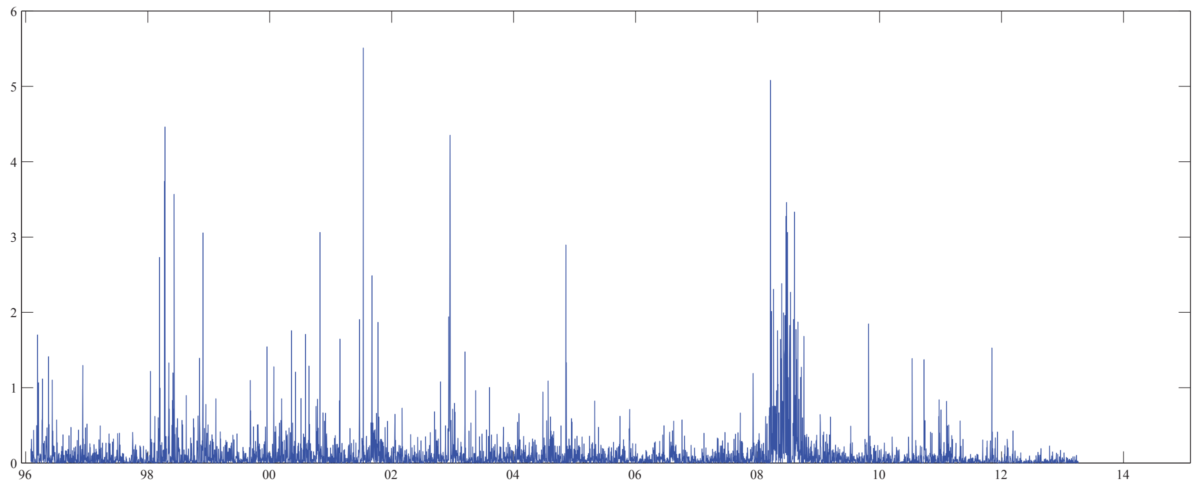


Figure 3: Daily squared return volatility proxy of the crude oil futures contracts maturing from 1 to 4 months.

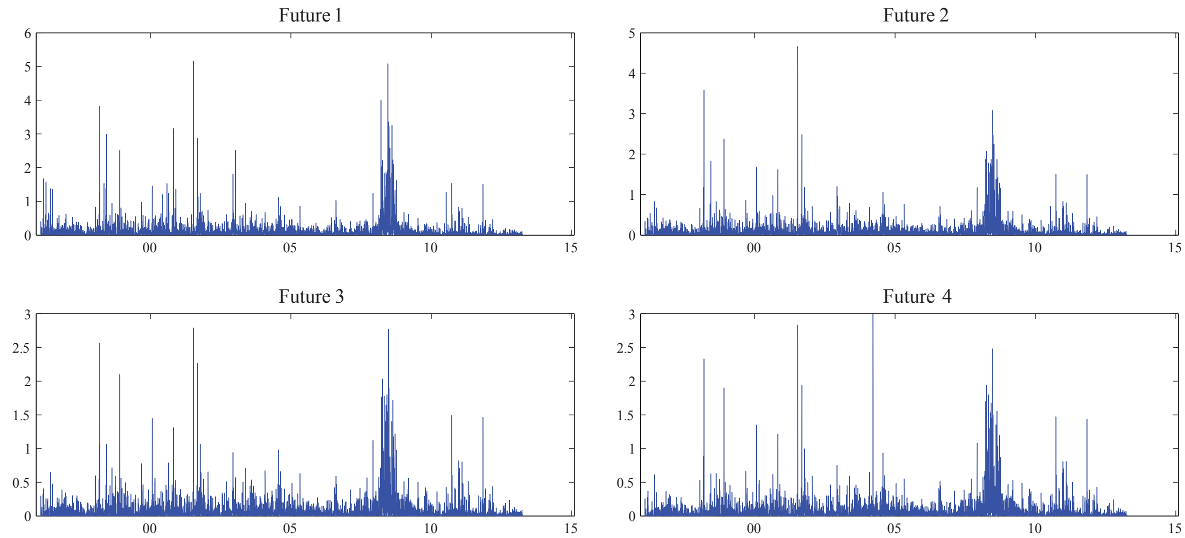


Table 4: Two-step Exact local Whittle estimates of δ for $m = \lfloor n^k \rfloor$ and $n = 4499$

k	Spot	Futures 1	Futures 2	Futures 3	Futures 4
0.4	0.3828	0.4473	0.4716	0.4852	0.4837
0.5	0.6037	0.6387	0.6643	0.6596	0.6357
0.6	0.3892	0.4155	0.4026	0.4267	0.4001
0.7	0.2585	0.3135	0.2962	0.3114	0.2947

bandwidth but heterogeneous across the bandwidths. Such variability is likely to appear in presence of level shifts or short run perturbations. Accordingly, in the following we shall apply the bandwidth requirement, $m = o(n^{2\delta/(1+2\delta)})$, suggested by Frederiksen et al. (2012) to prevent the presence of a neglected additive noise term in the volatility proxy. More precisely, we shall consider two bandwidths: $m = \lfloor n^k \rfloor$ for $k = \{0.4, 0.5\}$.

Table 5: W-test for $m = \lfloor n^k \rfloor$ and $\varepsilon = \{0.02, 0.05\}$

k	Spot Prices	Futures 1	Futures 2	Futures 3	Futures 4
$\varepsilon = 0.02$					
0.4	0.6261	0.6832	0.6664	0.6642	0.6600
0.5	0.8726	1.2638	0.8522	0.8303	0.7964
$\varepsilon = 0.05$					
0.4	0.6261	0.6832	0.6664	0.6642	0.6600
0.5	0.8726	1.2638	0.8522	0.8303	0.7964

Importantly, we also apply the procedure of Qu (2011) to test the null hypothesis that the volatility estimate is a stationary long memory process against the alternative of a process contaminated by level shifts or a smoothly varying trend. The so-called W-test statistic of Qu (2011) depends on a trimming parameter ε that we set to either 0.02 or 0.05 and for which there are two specific asymptotic critical value at a threshold of 10%: 1.118 and 1.022. We compute the test using the Gaussian semi-parametric estimator of Robinson (1995a). The results are reported in Table 5 and highlight that in all cases we do not reject the null of a stationary long memory process except for the volatility of the *Futures 1* for which it is difficult to conclude.

Table 6: Unbalance stationary cointegration analysis

k	0.4				0.5			
	Futures 1	Futures 2	Futures 3	Futures 4	Futures 1	Futures 2	Futures 3	Futures 4
T_0	3.98599	3.87855	3.36281	2.4622	1.22994	1.66711	1.06878	0.287712
$\hat{r}_{0.45}$	1	1	1	1	1	1	0	0
$\hat{r}_{0.35}$	1	1	1	1	1	1	1	1
$\hat{r}_{0.25}$	1	1	1	1	1	1	1	1
$\hat{\delta}_2$	0.4243	0.4621	0.4828	0.4858	0.652	0.6822	0.6807	0.6650
$\hat{\delta}_1$	0.1032	0.1646	0.2365	0.2989	0.0473	0.109	0.198	0.2326
$\hat{\xi}$	-0.0609	-0.097	-0.1171	-0.1208	-0.0414	-0.0756	-0.0775	-0.0614
$\hat{\beta}$	0.7094	0.4539	0.3767	0.3471	0.8423	0.5356	0.4966	0.5218
$\hat{v}$	0.3211	0.2975	0.2463	0.1869	0.6047	0.5732	0.4827	0.4324

We turn now to the core of our analysis that is the (non) equality of the integration orders. To investigate whether or not the integration orders of the pairwise volatilities are homogeneous, we apply the procedure of Nielsen and Shimotsu (2007). Because the cointegration is not observed, the authors propose to test for $H_0 : \delta_{F_i} = \delta_S$ with $i = 1, 2, 3, 4$ that is informative in both cases. However, computing the test statistic, $\hat{T}_0$, requires to estimate the cointegration rank, r and the authors suggest a model selection procedure based on a tuning parameter $v_k = m_G^{-\kappa} > 0$ where m_G is a specific bandwidth used to estimate

$\hat{G}$ and fixed to $m_G = \lfloor n^{k-0.05} \rfloor$. Here we consider $v_k = \{m_G^{-0.45}, m_G^{-0.35}, m_G^{-0.25}\}$ because the procedure in generally sensitive to the choice of κ. Nielsen and Shimotsu (2007) show that $\hat{T}_0 \xrightarrow{d} \chi_1^2$ if $r = 0$ and $\hat{T}_0 \xrightarrow{p} 0$ otherwise. The results are reported in the Table 6 and lead to ambiguous conclusions. Indeed, in several cases we accept the alternative hypothesis at a threshold of 10% (i.e. 2.71) although the rank estimates are positives in most cases.

In such inconclusive situation we support that unbalanced stationary cointegration is likely to occurs. The estimates $\hat{\beta}$, $\hat{\delta_2}$, $\hat{\delta_1}$ and $\hat{\xi}$ are also reported in Table 6. First of all, we observe that all $\hat{\xi}$ are negative and increases (in absolute value) with the maturity of the futures. According to the 2S-ELW estimates of δ_S and δ_{F_i}, this is not surprising. Also notice that the cointegration strengths, $\hat{v}$, are clearly non-null and decrease as the maturity of the futures increases. This result is very interesting with respect to hedging strategies and is consistent with the findings of Caporale et al. (2014). Moreover, it reveals that the no-arbitrage hypothesis seems valid for futures contracts maturing in one month whereas the results are more ambiguous for contracts with longer maturities. Indeed, the coefficient β also dramatically decreases as the maturity of the futures increases leading to lower long run hedging ratio for *Futures 2-4*. A possible explanation comes from the fact that the probability of occurrence of negative convenience yield - and hence the probability arbitrage opportunities - increases with the maturity.

7. Conclusion

In this paper we investigate the estimation of bivariate unbalanced stationary fractional cointegration models. The estimator we propose relies on the local behavior of the spectral density of the system in the vicinity of the origin. It allows for estimating jointly all parameters of interest of the model and notably the unbalanced parameter. We have demontrated the consistency of the estimator as well as its good finite sample properties estimator by means of Monte Carlo study. Our asymptotic results also suggest that neglecting for the presence of non-null unbalanced parameter might lead to inconsistency. In a short application we investigated the no-arbitrage hypothesis between the volatilities of spot and futures prices. Our results reveal that the apparent unbalance of the integration orders between the daily squared returns of the observable is misleading. An unbalanced stationary cointegration is recovered and the results are consistent with the theory as well as some empirical features found in the literature.

8. Appendix: Proof of theorem 1

Proof 1. *Let θ be the vector of admissible parameter value, θ_0 the vector of true parameter value and $S(\theta) = R_m(\theta) - R_m(\theta_0)$. Then, define the neighborhoods $\Theta_\delta^n = \{\delta : ||\delta - \delta_0|| < d\}$, $\Theta_\xi^n = \{\xi : ||\xi - \xi_0|| < e\}$, $\Theta_\beta^n = \{\beta : ||\lambda_m^{\delta_{01}-\delta_{02}}(\beta - \beta_0)|| < b\}$ and their complements $\Theta_\delta^c = \Theta_\delta \backslash \Theta_\delta^n$, $\Theta_\xi^c = \Theta_\xi \backslash \Theta_\xi^n$ and $\Theta_\beta^c = \Theta_\beta \backslash \Theta_\beta^n$ such that $\Theta = \Theta_\delta \times \Theta_\xi \times \Theta_\beta \backslash \Xi$. Without loss of generality with respect to Assumption 1 we set*

$$\max\left(\min_i ||\delta_i - \delta_{0i}||, ||\xi - \xi_0||\right) \geq d, \qquad \delta \in \Theta_\delta^c, \qquad \xi \in \Theta_\xi^c, \tag{20}$$

so that $1/2 > d \geq e > 0$. From Robinson (1995a, p. 1634) and by the fact that $\theta_0 \in \Theta_\delta^n \times \Theta_\xi^n \times \Theta_\beta^n$ it follows

$$\Pr\left(\{\hat{\delta} \in \Theta_\delta^c\} \cup \{\hat{\xi} \in \Theta_\xi^c\} \cup \{\hat{\beta} \in \Theta_\beta^c\}\right) =$$
$$\Pr\left(\inf_{\{\hat{\delta} \in \Theta_\delta^c\} \cup \{\hat{\xi} \in \Theta_\xi^c\} \cup \{\hat{\beta} \in \Theta_\beta^c\}} R_m(\theta) \leq \inf_{\{\hat{\delta} \in \Theta_\delta^n\} \cup \{\hat{\xi} \in \Theta_\xi^n\} \cup \{\hat{\beta} \in \Theta_\beta^n\}} R_m(\theta)\right)$$
$$\leq \Pr\left(\inf_{\{\hat{\delta} \in \Theta_\delta^c\} \cup \{\hat{\xi} \in \Theta_\xi^c\} \cup \{\hat{\beta} \in \Theta_\beta^c\}} S(\theta) \leq 0\right).$$

Accordingly, to prove the Theorem 1, it suffices to show that, as $n \to 0$, $S(\theta)$ is positive and bounded away from 0 uniformly on $\{\hat{\delta} \in \Theta_\delta^c\} \cup \{\hat{\xi} \in \Theta_\xi^c\} \cup \{\hat{\beta} \in \Theta_\beta^c\}$ so that

$$\Pr\left(\inf_{\{\hat{\delta} \in \Theta_\delta^c\} \cup \{\hat{\xi} \in \Theta_\xi^c\} \cup \{\hat{\beta} \in \Theta_\beta^c\}} S(\theta) \leq 0\right) \to 0. \tag{21}$$

Now, introduce $\psi_1 = \delta_1 - \delta_{01}$ and $\psi_2 = (\delta_2 + \xi) - (\delta_{02} + \xi_0)$ and develop $S(\theta)$ as

$$S(\theta) = \log\det\hat{G}(\theta) - 2(\delta_1 + \delta_2 + \xi) m^{-1} \sum_{j=1}^{m} \log\lambda_j$$
$$- \log\det\hat{G}(\theta_0) + 2(\delta_{01} + \delta_{02} + \xi_0) m^{-1} \sum_{j=1}^{m} \log\lambda_j$$
$$= \log\det\hat{G}(\theta) - \log\det\hat{G}(\theta_0) - m^{-1} \sum_{j=1}^{m} 2\log\lambda_j\,(\delta_1 + \delta_2 + \xi - \delta_{01} - \delta_{02} - \xi_0)$$
$$= \log\det\hat{G}(\theta) - \log\det\hat{G}(\theta_0) - m^{-1} \sum_{j=1}^{m} 2\log\lambda_j \sum_{i=1}^{p} \psi_i,$$

for $p = 2$ so that $\sum_{i=1}^{p} \psi_i = (\psi_1 + \psi_2)$. Then, by the fact that

$$\sum_{j=1}^{m} \log\lambda_j = \sum_{j=1}^{m} \log(2\pi j n^{-1})$$
$$= \sum_{j=1}^{m} \left(\log j + \log(2\pi n^{-1})\right)$$
$$= m\log(2\pi n^{-1}) + \sum_{j=1}^{m} \log j$$
$$m^{-1} \sum_{j=1}^{m} \log\lambda_j = \log(2\pi n^{-1}) + m^{-1} \sum_{j=1}^{m} \log j = \log\lambda_m + m^{-1} \sum_{j=1}^{m} \log j - \log m,$$

and rearranging $S(\theta)$, we obtain

$$\begin{aligned}S(\theta) &= \log\det\hat{G}(\theta) - \log\det\hat{G}(\theta_0) + \log\det G_0 - \log\det G_0 \\ &\quad - 2\log\lambda_m \sum_{i=1}^{p}\psi_i + 2\sum_{i=1}^{p}\psi_i\left(\log m - m^{-1}\sum_{j=1}^{m}\log j\right) \\ &\quad + \sum_{i=1}^{p}\log(2\psi_i+1) - \sum_{i=1}^{p}\log(2\psi_i+1),\end{aligned}$$

and finally, $S(\theta) = S_1(\theta) + S_2(\theta) + S_3(\theta)$, where

$$\begin{aligned}S_1(\theta) &= \log\det\hat{G}(\theta) - \log\det G_0 - 2\log\lambda_m \sum_{i=1}^{p}\psi_i + \sum_{i=1}^{p}\log(2\psi_i+1) \\ S_2(\theta) &= \log\det\hat{G}_0 - \log\det\hat{G}(\theta_0) \\ S_3(\theta) &= 2\sum_{i=1}^{p}\psi_i\left(\log m - m^{-1}\sum_{j=1}^{m}\log j\right) - \sum_{i=1}^{p}\log(2\psi_i+1).\end{aligned}$$

The way we split $S(\theta)$ has advantage that $S_2(\theta)$ and $S_3(\theta)$ do not depend on β so that we can treat them following the methodology of Robinson (1995a). To summarize, we have to demonstrate the boundedness of $S(\theta)$ under

$$\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^c\}\cup\{\hat{\xi}\in\Theta_\xi^c\}\cup\{\hat{\beta}\in\Theta_\beta^c\}} S(\theta) \le 0\right) \le$$

$$\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^c\}\cap\{\hat{\xi}\in\Theta_\xi^c\}\cap\{\hat{\beta}\in\Theta_\beta^c\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{22}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^c\}\cap\{\hat{\xi}\in\Theta_\xi^c\}\cap\{\hat{\beta}\in\Theta_\beta^n\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{23}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^c\}\cap\{\hat{\xi}\in\Theta_\xi^n\}\cap\{\hat{\beta}\in\Theta_\beta^n\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{24}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^n\}\cap\{\hat{\xi}\in\Theta_\xi^c\}\cap\{\hat{\beta}\in\Theta_\beta^n\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{25}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^c\}\cap\{\hat{\xi}\in\Theta_\xi^n\}\cap\{\hat{\beta}\in\Theta_\beta^c\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{26}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^n\}\cap\{\hat{\xi}\in\Theta_\xi^c\}\cap\{\hat{\beta}\in\Theta_\beta^c\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right) \tag{27}$$

$$+\Pr\left(\inf_{\{\hat{\delta}\in\Theta_\delta^n\}\cap\{\hat{\xi}\in\Theta_\xi^n\}\cap\{\hat{\beta}\in\Theta_\beta^c\}} \Big(S_1(\theta)+S_2(\theta)+S_3(\theta)\Big) \le 0\right). \tag{28}$$

We first consider $S_2(\theta)$ because it does not depend on $\hat{\theta}$ and no uniform bound is needed. Because $|\log(1+x)| \leq 2|x|$ (see Olver et al. 2010, p. 108), it follows that for $\varepsilon \leq 1$

$$\begin{aligned}\Pr\left(|S_2(\theta)| \leq \varepsilon\right) &= \Pr\left(|\log\det\hat{G}(\theta_0) - \log\det\hat{G}_0| \leq \varepsilon\right)\\ &\leq \Pr\left(\left|\frac{\det\hat{G}(\theta_0) - \det\hat{G}_0}{\det\hat{G}_0}\right| \leq \varepsilon/2\right).\end{aligned}$$

Accordingly, proving that $\det\hat{G}(\theta_0) - \det\hat{G}_0 \xrightarrow{p} 0$ suffices to show that $S_2(\theta)$ is $o_p(1)$ (see Robinson 1995a, p. 1635). To simplify notation, let $F^0_{jab} = G^0_{ab}\lambda_j^{-\delta_{0a}-\delta_{0b}}$, $f_z^{ab}(\lambda_j) = f^0_{jab}$, $A_{ab}(e^{i\lambda_j}) = A_{jab}$ and $I^0_{jab} = I^0_{ab}(\lambda_j;\vartheta_{02})$ with $\vartheta_{02} = (\beta_0,\xi_0)'$, G^0_{ab} the (a,b)-th element of G_0 and δ_{0a} is the ath element of $\vartheta_{01} = (\delta_{01},\delta_{02}+\xi_0)'$, respectively δ_{0b}. Evaluating Equations (6) and (8) at the true value, we obtain

$$I^0_j = \begin{pmatrix} I_{yy}(\lambda_j) - 2\beta_0\lambda_j^{\xi_0}I_{xy}(\lambda_j) + \beta_0^2\lambda_j^{2\xi_0}I_{xx}(\lambda_j) & I_{xy}(\lambda_j) - \beta_0\lambda_j^{\xi_0}I_{xx}(\lambda_j)\\ I_{xy}(\lambda_j) - \beta_0\lambda_j^{\xi_0}I_{xx}(\lambda_j) & I_{xx}(\lambda_j)\end{pmatrix},$$

from which we implicitly take the real part. Then,

$$\begin{aligned}\hat{G}_{ab}(\theta_0) - G^0_{ab} &= \frac{1}{m}\sum_{j=1}^{m}\frac{I^0_{jab}}{\lambda_j^{-\delta_{0a}-\delta_{0b}}} - G^0_{ab}\\ &= \frac{G^0_{ab}}{m}\sum_{j=1}^{m}\left(\frac{I^0_{jab}}{G^0_{ab}\lambda_j^{-\delta_{0a}-\delta_{0b}}} - 1\right) = \frac{G^0_{ab}}{m}\sum_{j=1}^{m}\left(\frac{I^0_{jab}}{F^0_{jab}} - 1\right),\end{aligned}$$

Recall that under Assumption 3 and 4, $f_z(\lambda_j) = A(e^{i\lambda_j})A(e^{i\lambda_j})^*/2\pi$ and $f_\varepsilon(\lambda_j) = I_2/2\pi$. Now, rewrite the latter expression as,

$$\begin{aligned}\frac{G^0_{ab}}{m}\sum_{j=1}^{m}\left(\frac{I^0_{jab}}{F^0_{jab}} - 1\right) &= S_{21}(\theta) + S_{22}(\theta) + S_{23}(\theta),\\ S_{21}(\theta) &= \frac{G^0_{ab}}{m}\sum_{j=1}^{m}\left(1 - \frac{F^0_{jab}}{f^0_{jab}}\right)\frac{I^0_{jab}}{F^0_{jab}},\\ S_{22}(\theta) &= \frac{G^0_{ab}}{m}\sum_{j=1}^{m}\frac{1}{f^0_{jab}}\left(I^0_{jab} - A_{jab}I_\varepsilon^{jab}\bar{A}_{jab})\right),\\ S_{23}(\theta) &= \frac{G^0_{ab}}{m}\sum_{j=1}^{m}\left(2\pi I_\varepsilon^{jab} - 1\right).\end{aligned}$$

From the analysis of (Robinson 1995a, p. 1636) and under Assumption 2-5, as $m \to \infty$, $|1 - F^0_{jab}(f^0_{jab})^{-1}| \leq \eta$, $E|I^0_{jab}(F^0_{jab})^{-1}| \leq c$ and thus $|S_{21}(\theta)| \leq c\eta$, with η and c any arbitrary positive numbers. Then next term to study *is $S_{22}(\theta)$. From Robinson (1995a, p. 1637), it can be demonstrated that*

$$E\left|I_{jab}^{0}-A_{jab}I_{\varepsilon}^{jab}\bar{A}_{jab}\right|=O\left(f_{jab}^{0}\log(j+1)^{1/2}j^{-1/2}\right),$$

and therefore $|S_{22}(\theta)|=o(1)$ as $m\to\infty$. Finally, as $n\to\infty$ and under Assumption 3,

$$\begin{aligned}
&2\pi I_{\varepsilon}-I_{2}=S_{231}(\theta)+S_{232}(\theta),\\
&\quad S_{231}(\theta)=n^{-1}\sum_{t=1}^{n}(\varepsilon_{t}\varepsilon_{t}'-I_{2})\xrightarrow{p}0,\\
&\quad S_{232}(\theta)=\sum_{s\neq t}^{n}\sum_{t}^{n}\cos\left((s-t)\lambda_{j}\right)\varepsilon_{s}\varepsilon_{t}=o(1),
\end{aligned}$$

and $S_{23}(\theta)$ is $o_{p}(1)$ (see Robinson 1995a, p. 1638). Thereby, as $n\to\infty$, we have proved that

$$\begin{aligned}
|\hat{G}_{ab}(\theta_{0})-G_{ab}^{0}|&=|S_{21}(\theta)|+|S_{22}(\theta)|+|S_{23}(\theta)|\\
&=O_{p}\left(\eta+m^{-1}\sum_{j=1}^{m}\left(\frac{\log j}{j}\right)^{1/2}\right)+o_{p}(1)=o_{p}(1),
\end{aligned}$$

and by corollary that $\det\hat{G}(\theta_{0})-\det\hat{G}_{0}\xrightarrow{p}0$. Straightforwardly, $S_{2}(\theta)=o_{p}(1)$ follows in Equations (22) to (28).

Now we turn to the analysis of $S_{3}(\theta)$. First, adding and subtracting $2(\psi_{1}+\psi_{2})$ to $S_{3}(\theta)$, we have

$$S_{3}(\theta)=-2\sum_{i=1}^{p}\psi_{i}\left(m^{-1}\sum_{j=1}^{m}\log j-\log m\right)-\sum_{i=1}^{p}\log(2\psi_{i}+1)+2\sum_{i=1}^{p}\psi_{i}-2\sum_{i=1}^{p}\psi_{i}$$

Then, rearranging these terms we obtain,

$$\begin{aligned}
S_{3}(\theta)=&-2\sum_{i=1}^{p}\psi_{i}\left(m^{-1}\sum_{j=1}^{m}\log j-(\log m-1)\right)\\
&+2\sum_{i=1}^{p}\psi_{i}-\sum_{i=1}^{p}\log(2\psi_{i}+1).
\end{aligned}\tag{29}$$

From Lemma 2 of Robinson (1995a) we know that

$$m^{-1}\sum_{j=1}^{m}\log j-(\log m-1)=O(m^{-1}\log m),$$

so that analysis of $S_{3}(\theta)$ reduces to study the greatest lower bound of (29) which is of the form $f(x)=x-\log(1+$

$x)$ for ψ_1 and $(x+y) - log(x+y+1)$ for ψ_2. Because $\inf_x f(x) \geq x^2/6$ and $\inf_{x,y} f(x,y) \geq (x^2+y^2)/6$ for $0 < |x| < 1$, $0 < |y| < 1$, and from the restriction stated in Equation (20), we can apply the analysis of Nielsen (2007, p. 437) uniformly over $\{\hat{\delta} \in \Theta_{\delta}^{c}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{c}\}$. From Lütkepohl (1996, sec. 8.5.2, p. 111) and by the triangular inequality,

$$\begin{aligned}&\sqrt{2}\max(|\delta_1 - \delta_{01}|, |\delta_2 - \delta_{02}|) + \sqrt{2}\max(|0|, |\xi - \xi_0|) \geq \\ &\left\| \begin{matrix} \delta_1 - \delta_{01} \\ \delta_2 - \delta_{02} \end{matrix} \right\| + \left\| \begin{matrix} 0 \\ \xi - \xi_0 \end{matrix} \right\| \geq ||\Psi|| \geq d + e,\end{aligned}$$

with $\Psi = (\psi_1, \psi_2)'$. Thereby, the infimum over $\{\hat{\delta} \in \Theta_{\delta}^{c}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{c}\} \cup \{\hat{\beta} \in \Theta_{\beta}\}$ of $2\sum_{i=1}^{p} \psi_i - \sum_{i=1}^{p} \log(2\psi_i + 1)$ is no less than

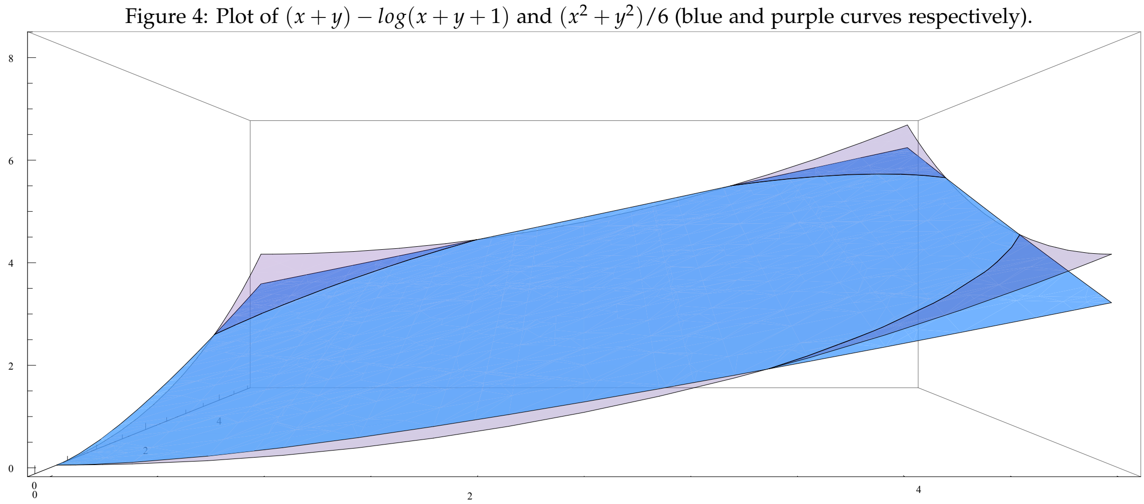
$$\frac{2(d+e)}{\sqrt{2}} - \log\left(1 + \frac{2(d+e)}{\sqrt{2}}\right) \geq \frac{2(d^2+e^2)}{6}.$$

Then, given that $f(x)$ has a unique minimum on $(-1, \infty)$ at $x = 0$

$$\begin{aligned}&\inf_{\{\hat{\delta} \in \Theta_{\delta}^{c}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{c}\} \cup \{\hat{\beta} \in \Theta_{\beta}\}} S_3(\theta) \geq \frac{2(d^2+e^2)}{6} + O(m^{-1}\log m), \\ &\inf_{\{\hat{\delta} \in \Theta_{\delta}^{n}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{n}\} \cup \{\hat{\beta} \in \Theta_{\beta}\}} S_3(\theta) = o(1).\end{aligned}$$

The two remaining cases are $\{\hat{\delta} \in \Theta_{\delta}^{c}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{n}\} \cup \{\hat{\beta} \in \Theta_{\beta}\}$ and $\{\hat{\delta} \in \Theta_{\delta}^{n}\} \cup \{\hat{\xi} \in \Theta_{\xi}^{c}\} \cup \{\hat{\beta} \in \Theta_{\beta}\}$. In the former case, $S_3(\theta)$ is no less than $2d^2/6 + O(m^{-1}\log m)$ while in the latter, $S_3(\theta)$ is no less than $2e^2/6 + O(m^{-1}\log m)$ (see graphic 4).

Figure 4: Plot of $(x+y) - log(x+y+1)$ and $(x^2+y^2)/6$ (blue and purple curves respectively).



Finally we turn to the analysis of $S_1(\theta)$. Rewrite $S_1(\theta)$ as

$$
\begin{aligned}
S_1(\theta) &= \log\det\hat{G}(\theta) - \log\det G_0 - 2\log\lambda_m \sum_{i=1}^{p}\psi_i + \sum_{i=1}^{p}\log(2\psi_i+1) \\
&= \log\det\hat{G}(\theta) - 2\log\lambda_m \\
&\times[\delta_1 - \delta_{01} + (\delta_2+\xi-\delta_{02}+\xi_0)] - \log\det G_0 + \sum_{i=1}^{p}\log(2\psi_i+1) \\
&= \log\det\hat{G}(\theta) - 2\log\lambda_m\psi_1 - 2\log\lambda_m\psi_2 - \log\left(\det G_0(2\psi_1+1)^{-1}(2\psi_2+1)^{-1}\right) \\
&= \log\det\left(V_m\hat{G}(\theta)V_m\right) - \log\left(\det G_0(2\psi_1+1)^{-1}(2\psi_2+1)^{-1}\right),
\end{aligned}
$$

where $\psi_2 = \delta_2 - \delta_{02} = (\delta_2+\xi) - (\delta_{02}+\xi_0)$, $\det G_0 = G_{11}G_{22} - G_{12}^2$, $V_m = \mathrm{diag}(\lambda_m^{-\psi_1}, \lambda_m^{-\psi_2})$ and

$$
\begin{aligned}
\det\left(V_m\hat{G}(\theta)V_m\right) &= \det\begin{pmatrix} \lambda_m^{-2\psi_1}\hat{G}_{11}(\theta) & \lambda_m^{-\psi_1-\psi_2}\hat{G}_{12}(\theta) \\ \lambda_m^{-\psi_1-\psi_2}\hat{G}_{21}(\theta) & \lambda_m^{-2\psi_2}\hat{G}_{22}(\theta) \end{pmatrix} \\
&= \lambda_m^{-2\psi_1-2\psi_2}\hat{G}_{11}(\theta)\hat{G}_{22}(\theta) - \lambda_m^{-2\psi_1-2\psi_2}\hat{G}_{21}(\theta)\hat{G}_{12}(\theta) \\
&= \lambda_m^{-2\psi_1-2\psi_2}(\hat{G}_{11}(\theta)\hat{G}_{22}(\theta) - \hat{G}_{12}^2(\theta)).
\end{aligned}
$$

Accordingly, analysis of $S_1(\theta)$ reduces to study

$$
S_1(\theta) = \lambda_m^{-2\psi_1-2\psi_2}\left(\hat{G}_{11}(\theta)\hat{G}_{22}(\theta) - \hat{G}_{12}^2(\theta)\right) - \left(G_{11}G_{22} - G_{12}^2\right)(2\psi_1+1)^{-1}(2\psi_2+1)^{-1}.
$$

Because $\hat{G}_{ab}(\theta) = m^{-1}\sum_{j=1}^{m}\lambda_j^{\delta_a+\delta_b}I_{jab}$,

$$
\begin{aligned}
S_1(\theta) &= S_{11}(\theta) + S_{12}(\theta) + S_{13}(\theta), \\
S_{11}(\theta) &= \lambda_m^{-2\psi_1-2\psi_2}\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta_1}I_{j11}\times m^{-1}\sum_{j=1}^{m}\lambda_j^{2(\delta_2+\xi)}I_{j22}\right), \\
S_{12}(\theta) &= -\lambda_m^{-2\psi_1-2\psi_2}\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{\delta_1+\delta_2+\xi}\,\mathrm{Re}(I_{j12})\right)^2, \\
S_{13}(\theta) &= -\left(G_{11}G_{22} - G_{12}^2\right)(1+2\psi_1)^{-1}(1+2\psi_2)^{-1}.
\end{aligned}
$$

Distinguishing the two summations by indexes j and k, multiplying by $G_{11}G_{22}/G_{11}G_{22}$ and $\lambda_j^{-2\delta_{01}}\lambda_j^{2\delta_{01}}$ and

$$
\lambda_k^{-2(\delta_{02}+\xi)}\lambda_k^{2(\delta_{02}+\xi_0)},
$$

and then rearranging $S_{11}(\theta)$ we have

$$
\begin{aligned}
S_{11}(\theta) &= \lambda_m^{-2\psi_1-2\psi_2}\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\lambda_j^{2\delta_1}\lambda_k^{2(\delta_2+\xi)}I_{j11}I_{k22}\lambda_j^{-2\delta_{01}}\lambda_j^{2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}\lambda_k^{2(\delta_{02}+\xi_0)} \\
&= \frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\frac{\lambda_j^{2\delta_1}\lambda_j^{-2\delta_{01}}\lambda_k^{2(\delta_2+\xi)}\lambda_k^{-2(\delta_{02}+\xi_0)}}{\lambda_m^{2\psi_1+2\psi_2}}\frac{I_{j11}I_{k22}}{\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}} \\
&= \frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{\lambda_j}{\lambda_m}\right)^{2\psi_1}\left(\frac{\lambda_k}{\lambda_m}\right)^{2\psi_2}\frac{I_{j11}I_{k22}}{\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}} \\
&= \frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1}\left(\frac{k}{m}\right)^{2\psi_2}\frac{I_{j11}I_{k22}}{\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}}.
\end{aligned}
$$

Now, rearranging $S_{12}(\theta)$ we obtain

$$
\begin{aligned}
S_{12}(\theta) &= -\lambda_m^{-2\psi_1-2\psi_2}\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{\delta_1+\delta_2+\xi}\operatorname{Re}(I_{j12})\right)^2 \\
&= -\lambda_m^{-2\psi_1-2\psi_2}\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\operatorname{Re}(I_{j12})\operatorname{Re}(I_{k12})\lambda_j^{\delta_1+\delta_2+\xi}\lambda_k^{\delta_1+\delta_2+\xi} \\
&= -\lambda_m^{-2\psi_1-2\psi_2}\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\operatorname{Re}(I_{j12})\operatorname{Re}(I_{k12}) \\
&\times \lambda_j^{\delta_1+\delta_2+\xi}\lambda_k^{\delta_1+\delta_2+\xi}\lambda_j^{\delta_{01}+\delta_{02}+\xi_0}\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{\delta_{01}+\delta_{02}+\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0} \\
&= -\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\frac{\lambda_j^{\delta_1+\delta_2+\xi}\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{\delta_1+\delta_2+\xi}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}}{\lambda_m^{2\psi_1+2\psi_2}}\frac{\operatorname{Re}(I_{j12})\operatorname{Re}(I_{k12})}{\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}} \\
&= -\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{\operatorname{Re}(I_{j12})\operatorname{Re}(I_{k12})}{\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}}.
\end{aligned}
$$

Then, we follow Nielsen (2007) and correct for the fact that $I_{11}(\lambda)$ and $I_{12}(\lambda)$ are based on estimated cointegration errors. Considering $I_{11}(\lambda) - I^0_{11}(\lambda)$ and $I_{12}(\lambda) - I^0_{12}(\lambda)$ we obtain

$$
\begin{aligned}
I_{11}(\lambda) &= I^0_{11}(\lambda) - 2\bar{\beta}\lambda_m^{\nu_0}\lambda^{\xi_0}\operatorname{Re}(I^0_{12}(\lambda)) + \bar{\beta}^2\lambda_m^{2\nu_0}\lambda^{2\xi_0}I_{22}(\lambda), \\
I_{12}(\lambda) &= I^0_{12}(\lambda) - \bar{\beta}\lambda_m^{\nu_0}\lambda^{\xi_0}I_{22}(\lambda).
\end{aligned}
$$

with $\bar{\beta}\lambda_m^{\nu_0} = (\beta - \beta_0)$ *and* $\nu_0 = \delta_{02} - \delta_{01}$*. Substituting in* $S_{11}(\theta)$ *and* $S_{12}(\theta)$ *we have*

$$
\begin{aligned}
S_{11}(\theta) &= \frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1}\left(\frac{k}{m}\right)^{2\psi_2}\frac{I_{j11}^0 I_{k22}}{\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}} \\
&- \frac{2\bar{\beta}}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1}\left(\frac{k}{m}\right)^{2\psi_2}\frac{\mathrm{Re}(I_{j12}^0) I_{k22}}{\lambda_m^{-\nu_0}\lambda_j^{-\xi_0}\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}}\frac{\lambda_j^{-\nu_0}}{\lambda_j^{-\nu_0}} \\
&+ \frac{\bar{\beta}^2}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1}\left(\frac{k}{m}\right)^{2\psi_2}\frac{I_{j22} I_{k22}}{\lambda_m^{-2\nu_0}\lambda_j^{-2\xi_0}\lambda_j^{-2\delta_{01}}\lambda_k^{-2(\delta_{02}+\xi_0)}}\frac{\lambda_j^{-2\nu_0}}{\lambda_j^{-2\nu_0}} \\
S_{12}(\theta) &= -\frac{1}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{\mathrm{Re}(I_{j12}^0)\,\mathrm{Re}(I_{k12}^0)}{\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}} \\
&+ \frac{2\bar{\beta}}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{I_{j22}\,\mathrm{Re}(I_{k12}^0)}{\lambda_m^{-\nu_0}\lambda_j^{-\xi_0}\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}}\frac{\lambda_j^{-\nu_0}}{\lambda_j^{-\nu_0}} \\
&- \frac{\bar{\beta}^2}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{I_{j22} I_{k22}}{\lambda_m^{-2\nu_0}\lambda_j^{-\xi_0}\lambda_j^{\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\xi_0}\lambda_k^{\delta_{01}-\delta_{02}-\xi_0}}\frac{\lambda_j^{-\nu_0}\lambda_k^{-\nu_0}}{\lambda_j^{-\nu_0}\lambda_k^{-\nu_0}},
\end{aligned}
$$

since $-I_{j12}I_{k12} = -I_{j12}^0 I_{k12}^0 + 2\bar{\beta}\lambda_m^{\nu_0}\lambda_j^{\xi_0} I_{j22} I_{k12}^0 - \bar{\beta}^2\lambda_m^{2\nu_0}\lambda_j^{\xi_0}\lambda_k^{\xi_0} I_{j22} I_{k22}$*. Rearranging this leads to*

$$
\begin{aligned}
S_{11}(\theta) &= \frac{G_{11}G_{22}}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1}\left(\frac{k}{m}\right)^{2\psi_2}\frac{I_{j11}^0 I_{k22}}{G_{11}\lambda_j^{-2\delta_{01}} G_{22}\lambda_k^{-2(\delta_{02}+\xi_0)}} \\
&\quad - 2\bar{\beta}\frac{G_{12}G_{22}}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1-\nu_0}\left(\frac{k}{m}\right)^{2\psi_2}\frac{\mathrm{Re}(I_{j12}^0) I_{k22}}{G_{12}\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0} G_{22}\lambda_k^{-2(\delta_{02}+\xi_0)}} \\
&\quad + \bar{\beta}^2\frac{G_{22}^2}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{2\psi_1-2\nu_0}\left(\frac{k}{m}\right)^{2\psi_2}\frac{I_{j22} I_{k22}}{G_{22}^2\lambda_j^{-2(\delta_{02}+\xi_0)}\lambda_k^{-2(\delta_{02}+\xi)}} \\
S_{12}(\theta) &= -\frac{G_{12}^2}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{\mathrm{Re}(I_{j12}^0)\,\mathrm{Re}(I_{k12}^0)}{G_{12}^2\lambda_j^{-\delta_{01}-\delta_{02}-\xi_0}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}} \\
&\quad + 2\bar{\beta}\frac{G_{12}G_{22}}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2-\nu_0}\left(\frac{k}{m}\right)^{\psi_1+\psi_2}\frac{I_{j22}\,\mathrm{Re}(I_{k12}^0)}{G_{12}\lambda_j^{-2(\delta_{02}+\xi_0)} G_{22}\lambda_k^{-\delta_{01}-\delta_{02}-\xi_0}} \\
&\quad - \bar{\beta}^2\frac{G_{22}^2}{m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{m}\right)^{\psi_1+\psi_2-\nu_0}\left(\frac{k}{m}\right)^{\psi_1+\psi_2-\nu_0}\frac{I_{j22} I_{k22}}{G_{22}^2\lambda_j^{-2(\delta_{02}-\xi_0)}\lambda_k^{-2(\delta_{02}-\xi_0)}},
\end{aligned}
$$

Now, consider the fact that $m^{-1}\sum_{j=1}^{m}(j/m)^{\alpha}=(1+\alpha)^{-1}$ *because*

$$m^{-1}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{\alpha}=\sum_{j=1}^{m}\int_{(j-1)/m}^{j/m}x^{\alpha}dx=\int_{0}^{1}x^{\alpha}dx,$$
$$\int_{0}^{1}x^{\alpha}dx=\left[\frac{1}{1+\alpha}x^{\alpha+1}\right]_{0}^{1}=(1+\alpha)^{-1},$$

Moreover, by the analysis of (Robinson 1995a, p. 1636-1638) we have

$$G_{ab}m^{-1}\sum_{j=1}^{m}\left(\frac{\mathrm{Re}(I_{jab}^{0})}{G_{ab}\lambda_{j}^{-\delta_{0a}-\delta_{0b}}}-1\right)=G_{ab}m^{-1}\left(\sum_{j=1}^{m}\mathrm{Re}(I_{jab}^{0})G_{ab}^{-1}\lambda_{j}^{\delta_{0a}+\delta_{0b}}-m\right)$$
$$=G_{ab}m^{-1}\sum_{j=1}^{m}\mathrm{Re}(I_{jab}^{0})G_{ab}^{-1}\lambda_{j}^{\delta_{0a}+\delta_{0b}}-1=o_{p}(1),$$

and thus it follows,

$$\begin{aligned}S_{11}(\theta)&=G_{11}G_{22}(1+2\psi_{1})^{-1}(1+2\psi_{2})^{-1}(1+o_{p}(1))\\&\quad-2\bar{\beta}G_{12}G_{22}(1+2\psi_{1}-\nu_{0})^{-1}(1+2\psi_{2})^{-1}(1+o_{p}(1))\\&\quad+\bar{\beta}^{2}G_{22}^{2}(1+2\psi_{1}-2\nu_{0})^{-1}(1+2\psi_{2})^{-1}(1+o_{p}(1)),\\S_{12}(\theta)&=-G_{12}^{2}(1+\psi_{1}+\psi_{2})^{-2}(1+o_{p}(1))\\&\quad+2\bar{\beta}G_{12}G_{22}(1+\psi_{1}+\psi_{2}-\nu_{0})^{-1}(1+\psi_{1}+\psi_{2})^{-1}(1+o_{p}(1))\\&\quad-\bar{\beta}^{2}G_{22}^{2}(1+\psi_{1}+\psi_{2}-\nu_{0})^{-2}(1+o_{p}(1)),\end{aligned}$$

Now we can rewrite $S_{1}(\theta)=S_{11}(\theta)+S_{12}(\theta)+S_{13}(\theta)$ *as*

$$\begin{aligned}S_{1}(\theta)&=G_{12}^{2}\left(\int_{0}^{1}x^{2\psi_{1}}dx\int_{0}^{1}x^{2\psi_{2}}dx-\left(\int_{0}^{1}x^{\psi_{1}+\psi_{2}}\right)^{2}\right)\\&\quad+\bar{\beta}^{2}G_{22}^{2}\left(\int_{0}^{1}x^{2\psi_{1}-2\nu_{0}}dx\int_{0}^{1}x^{2\psi_{2}}dx-\left(\int_{0}^{1}x^{\psi_{1}+\psi_{2}-\nu_{0}}\right)^{2}\right)\\&\quad+2\bar{\beta}G_{12}G_{22}\left(\int_{0}^{1}x^{\psi_{1}+\psi_{2}-\nu_{0}}dx\int_{0}^{1}x^{\psi_{1}+\psi_{2}}dx-\int_{0}^{1}x^{2\psi_{1}-\nu_{0}}dx\int_{0}^{1}x^{2\psi_{2}}dx\right)\\&\quad+G_{11}G_{22}\left((1+2\psi_{1})^{-1}(1+2\psi_{2})^{-1}-(1+2\psi_{1})^{-1}(1+2\psi_{2})^{-1}\right)+o_{p}(1).\end{aligned}$$

If $G_{12}=0$*, by the Cauchy-Schwarz inequality and the fact that* $\nu_{0}>0$ *under cointegration it is straightforward that* $S_{1}(\theta)\geq o_{p}(1)$ *when* $\{\hat{\delta}\in\Theta_{\delta}^{c}\}\cap\{\hat{\xi}\in\Theta_{\xi}^{c}\}\cap\{\hat{\beta}\in\Theta_{\beta}\}$ *and* $S_{1}(\theta)$ *is bounded away from zero when* $\{\hat{\delta}\in\Theta_{\delta}^{n}\}\cap\{\hat{\xi}\in\Theta_{\xi}^{n}\}\cap\{\hat{\beta}\in\Theta_{\beta}^{c}\}$*. A corollary of this finding is that* $S_{1}(\theta)$ *is also bounded when* $\{\hat{\delta}\in$

$\Theta^c_\delta\} \cap \{\hat{\xi} \in \Theta^n_\xi\} \cap \{\hat{\beta} \in \Theta^c_\beta\}$ and $\{\hat{\delta} \in \Theta^n_\delta\} \cap \{\hat{\xi} \in \Theta^c_\xi\} \cap \{\hat{\beta} \in \Theta^c_\beta\}$. Similarly, $S_1(\theta) \geq o_p(1)$ when $\{\hat{\delta} \in \Theta^n_\delta\} \cap \{\hat{\xi} \in \Theta^c_\xi\} \cap \{\hat{\beta} \in \Theta_\beta\}$ and $\{\hat{\delta} \in \Theta^c_\delta\} \cap \{\hat{\xi} \in \Theta^n_\xi\} \cap \{\hat{\beta} \in \Theta_\beta\}$. To summarize our results, for $c > 0$ an arbitrary small positive number and $B = 2(d^2 + e^2)/6 + o(1)$, we can rewrite the Equations (22) to (28) as

$$
\begin{aligned}
&\Pr\left(\inf_{\{\hat{\delta}\in\Theta^c_\delta\}\cup\{\hat{\xi}\in\Theta^c_\xi\}\cup\{\hat{\beta}\in\Theta^c_\beta\}} S(\theta) \leq 0\right) \\
&\leq \Pr\left(\inf_{\{\hat{\delta}\in\Theta^c_\delta\}\cap\{\hat{\xi}\in\Theta^c_\xi\}\cap\{\hat{\beta}\in\Theta^c_\beta\}} \Big([S_1(\theta) > c] + o_p(1) + [S_3(\theta) \geq B]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^c_\delta\}\cap\{\hat{\xi}\in\Theta^c_\xi\}\cap\{\hat{\beta}\in\Theta^n_\beta\}} \Big([S_1(\theta) \geq o_p(1)] + o_p(1) + [S_3(\theta) \geq B]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^c_\delta\}\cap\{\hat{\xi}\in\Theta^n_\xi\}\cap\{\hat{\beta}\in\Theta^n_\beta\}} \Big([S_1(\theta) \geq o_p(1)] + o_p(1) + [S_3(\theta) \geq \frac{2d^2}{6} + o(1)]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^n_\delta\}\cap\{\hat{\xi}\in\Theta^c_\xi\}\cap\{\hat{\beta}\in\Theta^n_\beta\}} \Big([S_1(\theta) \geq o_p(1)] + o_p(1) + [S_3(\theta) \geq \frac{2e^2}{6} + o(1)]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^c_\delta\}\cap\{\hat{\xi}\in\Theta^n_\xi\}\cap\{\hat{\beta}\in\Theta^c_\beta\}} \Big([S_1(\theta) > c] + o_p(1) + [S_3(\theta) \geq \frac{2d^2}{6} + o(1)]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^n_\delta\}\cap\{\hat{\xi}\in\Theta^c_\xi\}\cap\{\hat{\beta}\in\Theta^c_\beta\}} \Big([S_1(\theta) > c] + o_p(1) + [S_3(\theta) \geq \frac{2e^2}{6} + o(1)]\Big) \leq 0\right) \\
&+ \Pr\left(\inf_{\{\hat{\delta}\in\Theta^n_\delta\}\cap\{\hat{\xi}\in\Theta^n_\xi\}\cap\{\hat{\beta}\in\Theta^c_\beta\}} \Big([S_1(\theta) > c] + o_p(1) + [S_3(\theta) = o(1)]\Big) \leq 0\right),
\end{aligned}
$$

hence showing the Equation (21) and proving the Theorem 1.

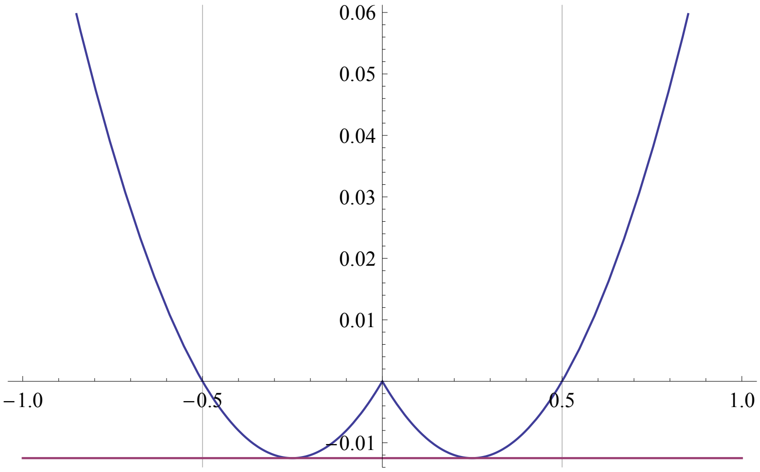
9. Extended Theorem 1

Now we extend the proof of the Theorem 1 to the case where $G_{12} = G_{21} \neq 0$. Given that G_{12} only impact $S_1(\theta)$ with regard to the positivity of $S(\theta)$ we only focus on the following expression

$$
\begin{aligned}
S_1(\theta) &= G_{12}^2\left(\int_0^1 x^{2\psi_1}dx \int_0^1 x^{2\psi_2}dx - \left(\int_0^1 x^{\psi_1+\psi_2}\right)^2\right) \\
&+ \bar{\beta}^2 G_{22}^2\left(\int_0^1 x^{2\psi_1-2\nu_0}dx \int_0^1 x^{2\psi_2}dx - \left(\int_0^1 x^{\psi_1+\psi_2-\nu_0}\right)^2\right) \\
&+ 2\bar{\beta} G_{12} G_{22}\left(\int_0^1 x^{\psi_1+\psi_2-\nu_0}dx \int_0^1 x^{\psi_1+\psi_2}dx - \int_0^1 x^{2\psi_1-\nu_0}dx \int_0^1 x^{2\psi_2}dx\right) + o_p(1), \\
S_1(\theta) &= S_{1a}(\theta) + S_{1b}(\theta) + S_{1c}(\theta) + o_p(1).
\end{aligned}
$$

Now observe that $S_{1a}(\theta) > 0$ by the Cauchy-Schwarz inequality. Then, we follow Nielsen (2007) and consider the fact that $S_{1c}(\theta) = 0$ when $\psi_1 = \psi_2$. In our unbalanced framework this is a very specific case because ψ_2 depends on ξ. Also recall that under cointegration, $\nu_0 > 0$. Accordingly, for $\eta > 0$ and $|\psi_2 - \psi_1| < \eta$, we have $S_{1b}(\theta) \geq \bar{\beta}^2 C$ with $C > 0$ and $|S_{1c}(\theta)| \leq |\bar{\beta}|\varepsilon$ with $\varepsilon > 0$ so that $-|\bar{\beta}|\varepsilon \leq S_{1c}(\theta) \leq |\bar{\beta}|\varepsilon$. Therefore, $S_1(\theta)$ is no less than $S_{1b}(\theta) - |\bar{\beta}|\varepsilon \geq \bar{\beta}^2 C - |\bar{\beta}|\varepsilon$.

Figure 5: Plot of $|\bar{\beta}|(|\bar{\beta}|C - \varepsilon)$, $|\bar{\beta}| \geq 2\varepsilon/C$ and $|\bar{\beta}| \leq 2\varepsilon/C$ (blue, black and purple curves respectively).



Setting to zero the first derivative of the bound we obtain that $S_1(\theta) > 0$ when $|\bar{\beta}| \geq 2\varepsilon/C$ and no less than $-\varepsilon^2/(4C)$ when $|\bar{\beta}| \leq 2\varepsilon/C$ (see Figure 5). Thus, for η sufficiently small, there exists ε such that 22 to 27 tend to 0 when $|\psi_2 - \psi_1| < \eta$.

Work in Progress

10. Appendix: Auxiliary result

In the following, we extend the Theorem 1 of Robinson and Marinucci (2003) and the Proposition 1 of Nielsen (2005) to the case of unbalanced cointegration. From Equation (8), for δ_1, δ_2 and ξ consistently pre-estimated, we have

$$\hat{G}_{ee}(\beta) = \hat{G}_{11}(\beta) = m^{-1} \sum_{j}^{m} \left(\lambda_j^{2\delta_1} I_{j11} \right) = m^{-1} \sum_{j}^{m} \left(\lambda_j^{2\delta_1} I_{jyy} + \beta^2 \lambda_j^{2\delta_1 + 2\xi} I_{jxx} - 2\beta \lambda_j^{2\delta_1 + \xi} \operatorname{Re}(I_{jxy}) \right),$$

with $e_t = y_t - \beta x_t$. Thus, $R_{m11}(\beta) = \log \hat{G}_{11}(\beta)$ and the derivative with respect to β is $\partial R_{m11}(\beta)/\partial \beta = 2m^{-1} \sum_j^m \left(\lambda_j^{2\delta_1 + \xi} \operatorname{Re}(\beta \lambda_j^{\xi} I_{jxx} - I_{jxy}) \right)$. Setting this equal to 0 we obtain

$$\hat{\beta} = \left(m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1 + \xi)} \operatorname{Re}(I_{jxx}) \right)^{-1} \left(m^{-1} \sum_{j=1}^{m} \lambda_j^{2(\delta_1 + \xi/2)} \operatorname{Re}(I_{jxy}) \right),$$

where $m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}\operatorname{Re}(I_{jab})$ is the (δ)-weighted periodogram and $\hat{\beta}$ the narrow-band generalized least squares (NBGLS) estimate of Nielsen (2005).

Proposition 1. *Under Assumptions 2, 3, 5 and for δ satisfying $(2(\delta_2+\xi)+2\delta_1-1)/4<\delta\leq\delta_1$,*

$$\hat{\beta}-\beta_0=\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{2(\delta_1+\xi)}\operatorname{Re}(I_{jxx})\right)^{-1}\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}\operatorname{Re}(I_{jxe})\right)=O_p\left((n/m)^{\delta_1-\delta_2}\right), \tag{30}$$

where $e_t=y_t-\beta x_t$.

Proof 2. *From the Proposition 1 of Nielsen (2005) we have*

$$\begin{aligned}m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}\operatorname{Re}(I_{jab})&=\int_0^{\lambda}\operatorname{Re}\left(f_{ab}(\mu)\right)d\mu\sim\frac{G_{ab}\lambda^{1-\delta_a-\delta_b+2\delta}}{1-\delta_a-\delta_b+2\delta}\\&=O_p\left(\lambda_m^{1-\delta_a-\delta_b+2\delta}\right).\end{aligned}$$

Therefore, by the Cauchy-Schwarz inequality

$$\begin{aligned}\left|m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta}\operatorname{Re}(I_{jxe})\right|&\leq\left(m^{-1}\sum_{j=1}^{m}\lambda_j^{2(\delta_1+\xi)}\operatorname{Re}(I_{jxx})\times m^{-1}\sum_{j=1}^{m}\lambda_j^{2\delta_1}\operatorname{Re}(I_{jee})\right)^{1/2}\\&\leq O_p\left(\left(\frac{m}{n}\right)^{1-2(\delta_2+\xi)+2(\delta_1+\xi)}\right)^{1/2}O_p\left(\left(\frac{m}{n}\right)^{1-2\delta_1+2\delta_1}\right)^{1/2}\\&\leq O_p\left(\left(\frac{m}{n}\right)^{1-\delta_2+\delta_1}\right).\end{aligned}$$

Finally, the proof is completed by substituting this result in Equation (30):

$$\begin{aligned}\hat{\beta}-\beta_0&=O_p\left(\left(\frac{m}{n}\right)^{-1+2(\delta_2+\xi)-2(\delta_1+\xi)}\right)O_p\left(\left(\frac{m}{n}\right)^{1-\delta_2+\delta_1}\right)\\&=O_p\left(\left(\frac{m}{n}\right)^{\delta_2-\delta_1}\right).\end{aligned}$$

11. Appendix: Simulation results

As specified in the simulation results section, there is weak evidence of consistency when the parameters lie in the non-stationary regions. Indeed, the RMSE decreases very slowly as the sample size increases. Mostly in the strong cointegration case, the bias increases as the sample size increases. Interestingly, results are less impacted by non-stationarity in the weak cointegration case. This suggests that practitioners should devote a particular attention to the stationarity or non-stationarity of the data. In the latter case, they should use the estimator of Hualde (2014).

Table 7: Simulation results for 10000 replications of the strong cointegration model when $\xi = 0.1$ and $\rho = 0$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.6	0	δ_2	0.024	0.027	0.166	0.032	0.016	0.131	0.032	0.009	0.101
		δ_1	0.169	0.045	0.271	0.197	0.027	0.257	0.234	0.017	0.267
		ξ	0.002	0.003	0.055	0.001	0.001	0.036	0.001	0.001	0.027
		β	0.039	0.013	0.122	0.043	0.010	0.110	0.050	0.008	0.103
0.8	0.2	δ_2	0.007	0.029	0.170	0.004	0.018	0.133	0.001	0.011	0.105
		δ_1	0.185	0.042	0.277	0.217	0.028	0.273	0.259	0.017	0.290
		ξ	-0.007	0.005	0.071	-0.007	0.002	0.048	-0.006	0.001	0.034
		β	-0.002	0.037	0.193	0.005	0.029	0.169	0.010	0.022	0.148
0.9	0.4	δ_2	-0.007	0.031	0.176	-0.007	0.021	0.144	-0.006	0.014	0.117
		δ_1	0.129	0.041	0.240	0.151	0.028	0.224	0.178	0.018	0.223
		ξ	-0.026	0.008	0.095	-0.024	0.004	0.069	-0.023	0.002	0.052
		β	-0.062	0.063	0.259	-0.070	0.051	0.236	-0.067	0.042	0.215
$m = \lfloor n^{0.8} \rfloor$											
0.6	0	δ_2	-0.048	0.004	0.080	-0.052	0.002	0.071	-0.063	0.001	0.072
		δ_1	0.085	0.005	0.113	0.118	0.004	0.133	0.154	0.003	0.163
		ξ	0.034	0.001	0.050	0.041	0.001	0.049	0.047	0.000	0.051
		β	0.104	0.006	0.129	0.149	0.005	0.166	0.205	0.005	0.217
0.8	0.2	δ_2	-0.051	0.005	0.087	-0.063	0.003	0.085	-0.083	0.002	0.096
		δ_1	0.092	0.006	0.122	0.133	0.005	0.151	0.180	0.005	0.192
		ξ	0.041	0.002	0.060	0.056	0.001	0.065	0.070	0.001	0.074
		β	0.104	0.015	0.160	0.183	0.013	0.216	0.287	0.012	0.308
0.9	0.4	δ_2	-0.057	0.006	0.095	-0.066	0.005	0.095	-0.086	0.004	0.105
		δ_1	0.067	0.006	0.104	0.089	0.005	0.115	0.120	0.005	0.139
		ξ	0.013	0.004	0.062	0.031	0.002	0.058	0.046	0.002	0.060
		β	0.027	0.036	0.191	0.092	0.036	0.212	0.182	0.037	0.265

References

Abadir, K., Talmain, G., 2002. Aggregation, Persistence and Volatility in a Macro Model. Review of Economic Studies 69, 749-779.

Abadir, K.M., Caggiano, G., Talmain, G., 2013. Nelson-Plosser revisited: The ACF approach. Journal of Econometrics 175, 22-34.

Baillie, R.T., Bollerslev, T., 2000. The forward premium anomaly is not as bad as you think. Journal of International Money and Finance 19, 471-488.

Brenner, R., Kroner, K., 1995. Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. Journal of Financial and Quantitative Analysis 30, 23-42.

Caporale, G., Ciferri, D., Girardi, A., 2014. Time-varying spot and futures oil price dynamics. Scottish Journal of Political Economy 61, 78-97.

Chakraborty, A., Evans, G.W., 2008. Can perpetual learning explain the forward-premium puzzle? Journal of Monetary Economics 55, 477-490.

Cheung, Y., Lai, K., 1993. A fractional cointegration analysis of purchasing power parity. Journal of Business & Economic Statistics 11, 103-112.

Chevillon, G., Mavroeidis, S., 2013. Learning can generate Long Memory. ESSEC Working Papers 1113.

Table 8: Simulation results for 10000 replications of the strong cointegration model when $\xi = 0.1$ and $\rho = 0.4$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.6	0	δ_2	0.068	0.027	0.178	0.067	0.015	0.141	0.057	0.009	0.111
		δ_1	0.157	0.039	0.252	0.176	0.022	0.229	0.209	0.013	0.239
		ξ	0.028	0.003	0.065	0.026	0.001	0.045	0.024	0.001	0.034
		β	0.175	0.019	0.222	0.183	0.014	0.217	0.191	0.010	0.217
0.8	0.2	δ_2	0.041	0.029	0.175	0.035	0.018	0.140	0.025	0.012	0.111
		δ_1	0.157	0.043	0.260	0.185	0.027	0.247	0.226	0.017	0.261
		ξ	0.006	0.006	0.075	0.007	0.002	0.050	0.008	0.001	0.035
		β	0.075	0.046	0.227	0.089	0.036	0.210	0.099	0.027	0.192
0.9	0.4	δ_2	0.010	0.033	0.181	0.009	0.021	0.145	0.004	0.015	0.121
		δ_1	0.111	0.046	0.241	0.125	0.030	0.213	0.152	0.019	0.204
		ξ	-0.014	0.008	0.093	-0.012	0.004	0.068	-0.009	0.002	0.050
		β	0.032	0.080	0.284	0.031	0.067	0.260	0.039	0.055	0.238
$m = \lfloor n^{0.8} \rfloor$											
0.6	0	δ_2	-0.067	0.004	0.091	-0.066	0.002	0.079	-0.068	0.001	0.074
		δ_1	0.108	0.005	0.130	0.134	0.004	0.148	0.159	0.003	0.167
		ξ	0.086	0.001	0.094	0.090	0.001	0.094	0.089	0.001	0.092
		β	0.295	0.007	0.306	0.361	0.006	0.369	0.433	0.005	0.439
0.8	0.2	δ_2	-0.068	0.005	0.098	-0.077	0.003	0.094	-0.093	0.002	0.102
		δ_1	0.116	0.007	0.143	0.156	0.005	0.172	0.200	0.005	0.212
		ξ	0.082	0.002	0.094	0.096	0.001	0.101	0.108	0.001	0.111
		β	0.255	0.018	0.289	0.367	0.016	0.388	0.506	0.015	0.520
0.9	0.4	δ_2	-0.077	0.006	0.108	-0.084	0.004	0.106	-0.102	0.003	0.116
		δ_1	0.082	0.007	0.118	0.104	0.006	0.128	0.135	0.005	0.152
		ξ	0.051	0.004	0.082	0.066	0.003	0.083	0.080	0.002	0.090
		β	0.191	0.050	0.294	0.279	0.050	0.358	0.399	0.051	0.459

Chevillon, G., Massmann, M., Mavroeidis, S., 2010. Inference in models with adaptive learning. Journal of Monetary Economics 57, 341-351.

Chow, Y., McAleer, M., Sequeira, J., 2000. Pricing of forward and futures contracts. Journal of Economic Surveys 14, 215-253.

Christensen, B.J., Nielsen, M.Ø., 2006. Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. Journal of Econometrics 133, 343-371.

Engle, R., Granger, C.W.J, 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica 55, 251-276.

de Truchis, G., 2013. Approximate Whittle Analysis of Fractional Cointegration and the Stock Market Synchronization Issue. Economic Modelling 24, 98-105.

Davidson, J., Hashimzade, N., 2009. Type I and type II fractional Brownian motions: A reconsideration. Computational Statistics & Data Analysis 53, 2089-2106.

Davidson, J., Sibbertsen, P., 2005. Generating schemes for long memory processes: regimes, aggregation and linearity. Journal of Econometrics 128, 253-282.

Dueker, M., Startz, R., 1998. Maximum-likelihood estimation of fractional cointegration with an application to US and Canadian bond rates. Review of Economics and Statistics 80, 420-426.

Frederiksen, P., Nielsen, F.S., Nielsen, M.Ø., 2012. Local polynomial Whittle estimation of perturbed fractional pro-

Table 9: Simulation results for 10000 replications of the weak cointegration model when $\xi = 0.1$ and $\rho = 0$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.6	0.4	δ_2	-0.017	0.034	0.185	-0.002	0.022	0.147	0.002	0.013	0.116
		δ_1	0.003	0.031	0.176	0.003	0.018	0.135	0.003	0.010	0.102
		ξ	0.013	0.015	0.123	0.006	0.007	0.086	0.004	0.004	0.062
		β	0.094	0.074	0.288	0.117	0.053	0.258	0.138	0.037	0.236
0.8	0.6	δ_2	-0.022	0.033	0.184	-0.010	0.021	0.144	-0.008	0.013	0.113
		δ_1	0.010	0.031	0.176	0.008	0.020	0.141	0.006	0.012	0.109
		ξ	0.020	0.021	0.146	0.013	0.011	0.107	0.012	0.006	0.079
		β	0.139	0.184	0.451	0.157	0.137	0.402	0.166	0.098	0.354
0.9	0.8	δ_2	-0.032	0.038	0.197	-0.025	0.027	0.165	-0.015	0.016	0.128
		δ_1	-0.000	0.029	0.171	-0.001	0.019	0.139	0.005	0.012	0.111
		ξ	0.009	0.033	0.182	0.004	0.022	0.148	-0.003	0.012	0.109
		β	0.107	0.388	0.632	0.083	0.271	0.527	0.064	0.196	0.447
$m = \lfloor n^{0.8} \rfloor$											
0.6	0.4	δ_2	-0.052	0.004	0.085	-0.048	0.002	0.067	-0.047	0.001	0.059
		δ_1	-0.013	0.004	0.061	-0.007	0.002	0.044	-0.002	0.001	0.032
		ξ	0.035	0.003	0.065	0.037	0.001	0.054	0.038	0.001	0.048
		β	0.179	0.016	0.218	0.203	0.012	0.230	0.224	0.010	0.245
0.8	0.6	δ_2	-0.061	0.004	0.090	-0.059	0.002	0.077	-0.062	0.001	0.073
		δ_1	-0.003	0.004	0.061	0.007	0.002	0.047	0.014	0.001	0.040
		ξ	0.029	0.003	0.062	0.036	0.002	0.054	0.043	0.001	0.053
		β	0.239	0.042	0.314	0.281	0.033	0.334	0.332	0.028	0.372
0.9	0.8	δ_2	-0.046	0.006	0.090	-0.036	0.003	0.066	-0.031	0.002	0.053
		δ_1	-0.014	0.004	0.064	-0.005	0.002	0.049	0.002	0.001	0.038
		ξ	-0.014	0.005	0.069	-0.010	0.002	0.050	-0.005	0.001	0.038
		β	0.058	0.093	0.311	0.072	0.078	0.288	0.091	0.073	0.286

cesses. Journal of Econometrics 167, 426-447.

Gil-Alana, L.A., Hualde, J., 2009. Fractional integration and cointegration. An overview and an empirical application. Palgrave handbook of Econometrics, Vol. 2. Applied Econometrics, K. Patterson and T.C. Mills eds, Palgrave, MacMillan, Vol. 2, 434-469.

Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227-238.

Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specification. Journal of econometrics 16, 121-130.

Granger, C.W.J., 1986. Developments in the study of cointegrated economic variables. Oxford Bulletin of Economics and Statistics 4, 221-238.

Granger, C.W.J., 2010. Some thoughts on the development of cointegration. Journal of Econometrics 158, 3-6.

Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70, 393-408.

Hassler, U., Marmol, F., Velasco, C., 2006. Residual log-periodogram inference for long-run relationships. Journal of Econometrics 130, 165-207.

Helgason, H., Pipiras, V., Abry, P., 2011. Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding. Signal Processing 91, 1123-1133.

Table 10: Simulation results for 10000 replications of the weak cointegration model when $\xi = 0.1$ and $\rho = 0.4$

$m = \lfloor n^{0.5} \rfloor$			256			512			1024		
δ_2	δ_1	$\hat{\theta}$	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
0.6	0.4	δ_2	-0.007	0.034	0.185	0.009	0.022	0.147	0.014	0.013	0.114
		δ_1	-0.003	0.032	0.179	-0.011	0.020	0.141	-0.011	0.011	0.107
		ξ	0.030	0.011	0.111	0.022	0.005	0.076	0.019	0.002	0.053
		β	0.413	0.091	0.511	0.434	0.064	0.502	0.455	0.043	0.500
0.8	0.6	δ_2	-0.019	0.034	0.184	-0.008	0.022	0.149	-0.001	0.013	0.113
		δ_1	0.002	0.033	0.182	-0.000	0.021	0.145	-0.005	0.012	0.109
		ξ	0.029	0.016	0.128	0.024	0.008	0.094	0.021	0.004	0.068
		β	0.407	0.221	0.622	0.421	0.158	0.579	0.435	0.115	0.552
0.9	0.8	δ_2	-0.037	0.039	0.200	-0.025	0.024	0.158	-0.015	0.017	0.131
		δ_1	-0.001	0.032	0.180	-0.007	0.021	0.143	-0.005	0.013	0.116
		ξ	0.012	0.027	0.166	0.005	0.015	0.123	0.000	0.009	0.094
		β	0.374	0.454	0.771	0.350	0.325	0.669	0.336	0.225	0.581
$m = \lfloor n^{0.8} \rfloor$											
0.6	0.4	δ_2	-0.064	0.004	0.088	-0.057	0.002	0.071	-0.053	0.001	0.061
		δ_1	-0.007	0.003	0.056	-0.005	0.002	0.041	-0.004	0.001	0.031
		ξ	0.065	0.002	0.081	0.064	0.001	0.073	0.061	0.001	0.067
		β	0.502	0.017	0.518	0.525	0.012	0.536	0.544	0.010	0.552
0.8	0.6	δ_2	-0.071	0.004	0.094	-0.065	0.002	0.079	-0.064	0.001	0.073
		δ_1	-0.004	0.004	0.059	0.003	0.002	0.045	0.007	0.001	0.038
		ξ	0.059	0.003	0.079	0.064	0.001	0.074	0.068	0.001	0.074
		β	0.523	0.047	0.567	0.569	0.036	0.599	0.623	0.030	0.646
0.9	0.8	δ_2	-0.053	0.005	0.089	-0.042	0.003	0.068	-0.035	0.002	0.053
		δ_1	-0.015	0.004	0.066	-0.007	0.003	0.051	-0.003	0.002	0.039
		ξ	-0.001	0.004	0.061	0.002	0.002	0.046	0.006	0.001	0.036
		β	0.339	0.104	0.468	0.342	0.088	0.452	0.365	0.081	0.462

Hualde, J., 2006. Unbalanced Cointegration. Econometric Theory 22, 765-814.

Hualde, J., 2012. Weak convergence to a modified fractional Brownian motion. Journal of Time Series Analysis 33, 519-529.

Hualde, J., 2013. A simple test for the equality of integration orders. Economics Letters 119, 233-237.

Hualde, J., 2014. Estimation of long-run parameters in unbalanced cointegration. Journal of Econometrics 178, 761-778.

Hualde, J., Robinson, P. M., 2007. Root-n-consistent estimation of weak fractional cointegration. Journal of Econometrics, 140, 450-484.

Hualde, J., Robinson, P.M., 2010. Semiparametric inference in multivariate fractionally cointegrated systems. Journal of Econometrics 157, 492-511.

Hurvich, C.M., Ray, B.K., 1995. Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. Journal of Time Series Analysis, 16, 17-42.

Johansen, S., Nielsen, M.Ø., 2012. Likelihood Inference for a Fractionally Cointegrated Vector Autoregressive Model. Econometrica 80, 2667-2732.

Liu, P., Tang, K., 2010. No-arbitrage conditions for storable commodities and the modeling of futures term structures. Journal of Banking & Finance 34, 1675-1687.

Liu, P., Tang, K., 2011. The stochastic behavior of commodity prices with heteroskedasticity in the convenience yield. Journal of Empirical Finance 18, 211-224.

Lobato, I.N., 1999. A semiparametric two-step estimator in a multivariate long memory model. Journal of Econometrics 90, 129-153.

Lütkepohl, H., 1996. Handbook of Matrices, New York: Wiley.

Marinucci, D., Robinson, P.M., 1999. Alternative forms of fractional Brownian motion. Journal of Statistical Planning and Inference 80, 111-122.

Maynard, A., Phillips, P., 2001. Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. Journal of Applied Econometrics 16, 671-708.

Miller, J.I., Park, J.Y., 2010. Nonlinearity, nonstationarity, and thick tails: How they interact to generate persistence in memory. Journal of Econometrics 155, 83-89.

Nielsen, M.Ø., 2004. Optimal Residual-Based Tests for Fractional Cointegration and Exchange Rate Dynamics. Journal of Business & Economic Statistics 22, 331-345.

Nielsen, M.Ø., 2005. Semiparametric Estimation in Time-Series Regression with Long-Range Dependence. Journal of Time Series Analysis 26, 279-304.

Nielsen, M.Ø., 2007. Local Whittle Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation. Journal of Business & Economic Statistics 25, 427-446.

Nielsen, M.Ø., Frederiksen, P., 2011. Fully modified narrow-band least squares estimation of weak fractional cointegration. The Econometrics Journal 14, 77-120.

Nielsen, M.Ø., Shimotsu, K., 2007. Determining the cointegrating rank in nonstationary fractional systems by the exact local Whittle approach. Journal of Econometrics 141, 574-596.

Olver, F., Lozier, D., Boisvert, R., 2010. NIST Handbook of Mathematical Functions. Cambridge University Press.

Perron, P., Qu, Z., 2007. An analytical evaluation of the log-periodogram estimate in the presence of level shifts. Boston University Working paper series.

Phillips, P., 1991. Optimal inference in cointegrated systems. Econometrica 59, 283-306.

Qu, Z., 2011. A test against spurious long memory. Journal of Business & Economic Statistics 29, 423-438.

Robinson, P., 1978. Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics 5, 163-168.

Robinson, P.M., 1994. Semiparametric analysis of long-memory time series. The Annals of Statistics 22, 515-539.

Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence. The Annals of statistics 23, 1630-1661.

Robinson, P.M., 1995. Log-periodogram regression of time series with long range dependence. The Annals of Statistics 23, 1048-1072.

Robinson, P.M., 2005. The distance between rival nonstationary fractional processes. Journal of Econometrics 128, 283-300.

Robinson, P., 2008. Multiple local Whittle estimation in stationary systems. The Annals of Statistics 36, 2508-2530.

Robinson, P., Henry, M., 1999. Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. Econometric theory 15, 299-336.

Robinson, P.M., Hualde, J., 2003. Cointegration in fractional systems with unknown integration orders. Econometrica 71, 1727-1766.

Robinson, P.M., Marinucci, D., 2003. Semiparametric frequency domain analysis of fractional cointegration. In Time Series with Long Memory (ed. P. M. Robinson). Oxford University Press, pp. 334-373.

Robinson, P.M., Yajima, Y., 2002. Determination of cointegrating rank in fractional systems. Journal of Econometrics 106, 217-241.

Rossi, E., Santucci de Magistris, P., 2013. A no-arbitrage fractional cointegration model for futures and spot daily ranges. Journal of Futures Markets 33, 77-102.

Schennach, S., 2013. Long memory via networking. CEMMAP Working Paper Nº 13.

Shimotsu, K., 2007. Gaussian semiparametric estimation of multivariate fractionally integrated processes. Journal of Econometrics 137, 277-310.

Shimotsu, K., 2010. Exact local Whittle estimation of fractional integration with unknown mean and time trend. Econometric Theory 1-43.

Shimotsu, K., 2012. Exact local Whittle estimation of fractionally cointegrated systems. Journal of Econometrics 169, 266-278.

Thornton, M.A., 2014. The aggregation of dynamic relationships caused by incomplete information. Journal of Econometrics 178, 342-351.

Tsay, W. J., Chung, C. F., 2000. The spurious regression of fractionally integrated processes. Journal of Econometrics, 96, 155-182.

Velasco, C., 1999a. Non-stationary log-periodogram regression. Journal of Econometrics 91, 325-371.

Velasco, C., 1999b. Gaussian semiparametric estimation for non-stationary time series. Journal of Time Series Analysis, 20, 87-127.

Velasco, C., 2003. Gaussian semi-parametric estimation of fractional cointegration. Journal of Time Series Analysis, 24, 345-378.

Zaffaroni, P., 2004. Contemporaneous aggregation of linear dynamic models in large economies. Journal of Econometrics 120, 75-102.

Zaffaroni, P., 2007. Aggregation and memory of models of changing volatility. Journal of Econometrics 136, 237-249.